

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:19:57 ; Search time 23 Seconds
(without alignments)
1087.280 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRLGKILSVIGLLFVL.....BEMAKELVELSKDKKDK 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	24.0	330	4	US-09-830-217-16
2	405	24.0	330	4	US-08-956-171B-5193
3	405	24.0	330	4	US-08-781-986A-5193
4	405	24.0	330	4	US-10-278-946-16
5	307.5	18.2	352	4	US-09-107-532A-4516
6	296.5	17.5	351	3	US-09-134-001C-5492
7	296	17.5	351	4	US-09-543-681A-6862
8	292	17.3	347	4	US-09-710-279-462
9	292	17.3	347	4	US-09-710-279-1268
10	283	16.7	317	4	US-09-071-035-38
11	283	16.7	327	4	US-09-134-000C-4342
12	278.5	16.5	325	4	US-09-543-681A-5565
13	268	15.9	296	4	US-09-071-035-40
14	266	15.7	306	4	US-09-602-787A-572
15	265	15.7	60	3	US-09-134-001C-4860
16	254	15.0	344	4	US-09-583-110-3723
17	250	14.8	321	4	US-09-583-110-2973
18	245.5	14.6	332	4	US-09-107-433-4809
19	241	14.3	324	4	US-09-602-787A-598
20	240	14.2	312	4	US-09-328-352-4444
21	240	14.0	324	4	US-09-071-035-70
22	236	14.0	298	3	US-08-961-083-24
23	236	14.0	298	4	US-09-536-784-24
24	236	14.0	301	4	US-09-134-000C-6014
25	228	13.5	289	4	US-09-071-035-72
26	218	12.9	321	4	US-09-302-626B-12
27	217	12.8	321	4	US-09-302-626B-10

ALIGNMENTS

RESULT 1

US-09-830-217-16
; Sequence 16, Application US/09830217
; Patent No. 6521441
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461PCT
; CURRENT APPLICATION NUMBER: US/09/830,217
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-830-217-16

Query Match	24.0%	Score 405;	DB 4;	Length 330;
Best Local Similarity	32.8%	Pred. No. 1.9e-25;		
Matches 113;	Conservative 65;	Mismatches 132;	Indels 34;	Gaps 12;
QY	8	LKILSVIGLLFVLIATAACGNSSNSSKSSKOGVEIKHEGTTKVPKPKRVVTVLEYS	67	
Db	5	IKML-VVTALFLV-LACGSGNSNQSDNDKDKETTSIKHANGTTEIKGPKRVTVLYG	62	
QY	68	FVDALVALDVPVGIADNDKKNRIILPLRDKIGKTVTSVTRKQPNLEISIKLPDIIAD	127	
Db	63	ATDVAVSLGVKPVGAVESWTQPKPEYIKNDLKDTPKIVGQEPAPNLEISIKLPDIIAS	122	
QY	128	NNRHGIVKDKLNIAPTIELKSFDDYDYNENIDAFKTSKALGKEEGKRLLEHDKKIEE	187	
Db	123	KVRNEKVDQLSKIAPTV---STDVTFKFK-DTTLGMLGKALGKEAEADLLKKYDDKVA	178	
QY	198	YKKEITMDKNQKVLPAVAAGSLLAHPSN-----SYVGQFSLSGPKEALSDDVTGKL	240	
Db	179	FOK----DAKAKYKDAWPLKASVNVFRADHTRIYAGGATAGELNDLGFKR-----	228	
QY	241	SKYLKG--PYLQMT--ETLSQVNPFRMTMTNKASNEPSL-----KELEKDPVVKKLA	292	
Db	229	QKQVNGKDIILTSTKESIPLMNADHI FVVKSDPNKDAALVKKTESEWTSKWKQLDA	288	
QY	293	VKNQRV-DILDRDLWAPRGLISSEMAKELVELSKDKKDK 335		

Db 289 VKNQVSDDLDEITWNLGGYKSSLSKLDDLYE--KLNIEKQSK 330

RESULT 2

US-08-956-171E-5193

Sequence 5193, Application US/08956171E

Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 5193:

SEQUENCE CHARACTERISTICS:

LENGTH: 330 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 5193:

US-08-956-171E-5193

Query Match 24.0%; Score 405; DB 4; Length 330;

Best Local Similarity 32.8%; Pred. No. 1.9e-25;

Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

Qy 8 LKILSVIGLLFVLIIATAACGNSSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLEYS 67

Db 5 IKML-VVTLAFLV-LAGCSGNSKQSSDNKDKEITTSIKHANGTTEIKGPKRVVTLVQG 62

Qy 68 FVDALVALDVKPGVIADNKNRIIKPLRDKIGKTSVGTGRQPNLEISKLKPDLIAD 127

Db 63 ATDVAVSLGVKPGVAVESWTQKPKFYIKNDLKDTKIVGQSPAPNLEISKLKPDLIAS 122

Qy 128 NNRHKGIVKDLNKTAPTIETLSKDFGQDYNENIDAFKTSIKALGKEEGKRLKEEHDKITEE 187

Db 123 KVRNEKYDQLSKLAPTV---STDVTFKFK-DTTKMGKALGKEKEAEDLLKKYDDKVA 178

Qy 188 YKKEITMDKNQKVLPAVAASKGLLAHPSN-----SYVGQFLSQLGFKSEALSDVTKGL 240

Db 179 FQK----DAKAKYKDAWPLKASVNVFRADHTRIYAGGYAGEILNDLGFKR-----NKDL 228

Qy 241 SKYLGK--PYLQWMT-ETLSQVNPMPIMTNKASSNEPSL-----KELEKDPVWKKLN 292

Db 229 QKQVDNGKDIQILTSKESIPLMNADHIFVVKSDPNADKDALVVKTESEWTSKWKKNLDA 288

Qy 293 VKNQRV-DILDRDLWARSRLISSEEMAKELVELSKDKSKDNK 335

Db 289 VKNQVSDDLDEITWNLGGYKSSLSKLDDLYE--KLNIEKQSK 330

RESULT 3

US-08-781-986A-5193

Sequence 5193, Application US/08781986A

Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 5193:

SEQUENCE CHARACTERISTICS:

LENGTH: 330 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-781-986A-5193

Query Match 24.0%; Score 405; DB 4; Length 330;

Best Local Similarity 32.8%; Pred. No. 1.9e-25;

Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

Qy 8 LKILSVIGLLFVLIIATAACGNSSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLEYS 67

Db 5 IKML-VVTLAFLV-LAGCSGNSKQSSDNKDKEITTSIKHANGTTEIKGPKRVVTLVQG 62

Qy 68 FVDALVALDVKPGVIADNKNRIIKPLRDKIGKTSVGTGRQPNLEISKLKPDLIAD 127

Db 63 ATDVAVSLGVKPGVAVESWTQKPKFYIKNDLKDTKIVGQSPAPNLEISKLKPDLIAS 122

Qy 128 NNRHKGIVKDLNKTAPTIETLSKDFGQDYNENIDAFKTSIKALGKEEGKRLKEEHDKITEE 187

Db 123 KVRNEKYDQLSKLAPTV---STDVTFKFK-DTTKMGKALGKEKEAEDLLKKYDDKVA 178

Qy 188 YKKEITMDKNQKVLPAVAASKGLLAHPSN-----SYVGQFLSQLGFKSEALSDVTKGL 240

Db 179 FQK----DAKAKYKDAWPLKASVNVFRADHTRIYAGGYAGEILNDLGFKR-----NKDL 228

Db 183 ENLGKIYDKEDKAKINKDLDRKISDMKD-KTKDFNKKVMYLLVNEGESTFGPGGRFGG 241

QY 221 QFSLGPFKEALSDVTKGLSKYLGYPYLO-MNTETLSQVNP-----ERFMINTKASS 273
DB 242 LVFDTLGFKA-----DKKVSXSPHGQINNEYINKQNPDPVILAMDRGVSVGGKATT 293
QY 274 NEPSLKELEKDPVWKLANVKNQVDILDRDLWARSRLISSEEMAKELVELSK 327
DB 294 N-----QVLKKNVKNVKAWSNHIYELDPKLYFSSG--SSTTTIKQIDELNE 340

RESULT 9

US-09-710-279-1268
; Sequence 1268, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1268
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1268

Query Match 17.3%; Score 292; DB 4; Length 347;
Best Local Similarity 28.5%; Pred. No. 4.1e-16;
Matches 101; Conservative 68; Mismatches 121; Indels 64; Gaps 16;

QY 16 LFLVLIATAACGNSSSSSKE-----SSKDGVEIKHE-----EGTKVP 55
DB 9 LLSLVLIATAACGNSSSSSKE-----SSKDGVEIKHE-----EGTKVP 55
QY 56 KHPKRVVLEYSFVDALVDKPGIADNKK-----KNRIIKPLRD--KIGKTSVGT 107
DB 69 KNPKNVLDYDGLDVL-----KELGVAUKVGLPKGKNNQSLPKFLDFDKDKYINTGN 123
QY 108 RKOPNLEESKLPDLI--IADNNRHGIYKDLNKAAPTIELKSPDGDYNNENI--DAFK-- 162
DB 124 LKEVNPDKVASAKPDVIFISGRANQNLDEFKKAAPKAKV-VYVGTSDNLIKDKMKNT 182
QY 163 -TISKALGEEGKKELEHDKIEYKKEITMDKNQKVLPAVAASKGLLAH-PSNSVYG 220
DB 183 ENLGKIYDKEDAKKINKOLDRIKSDMKDK-TKDFNKKVYLLVNEGELSTFGPGRFG 241
QY 221 QFSLGPFKEALSDVTKGLSKYLGYPYLO-MNTETLSQVNP-----ERFMINTKASS 273
DB 242 LVFDTLGFKA-----DKKVSXSPHGQINNEYINKQNPDPVILAMDRGVSVGGKATT 293
QY 274 NEPSLKELEKDPVWKLANVKNQVDILDRDLWARSRLISSEEMAKELVELSK 327
DB 294 N-----QVLKKNVKNVKAWSNHIYELDPKLYFSSG--SSTTTIKQIDELNE 340

RESULT 10

US-09-071-035-38
; Sequence 38, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue

; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-38

Query Match 16.7%; Score 283; DB 4; Length 317;

Best Local Similarity 30.4%; Pred. No. 2e-15;
Matches 101; Conservative 59; Mismatches 130; Indels 42; Gaps 13;
QY 11 LSVIGLFLVLIATAACGN-----SSNSKSSKSGKGVKHEEGTKVPHKRVVVLAYS 67
DB 11 VSMGLML-----SACQTKKTADSATTETAKTEVTVDKTNGLTVPKPKVVFONG 66
QY 68 FVDALVALDV--KPGIADNKNRIIKPLRDKIGKTSVGTROPNLEESKLPDLII 125
DB 67 SLDTMDALGVDRVVGAPTKN-----IPAYLKVKQKVESAGIKPEPDLKINQLKPDII 121
QY 126 ADNNRHGIYKDLNKAAPTIEL-----KSGFDGYNENIDAFKTIKALGKEEGKRLBEH 181
DB 122 I-SGQQDYQQLKAIAPTIYLVADKNPWASTKQNIETLGTI---FDKEEVAKEKITGL 177
QY 182 DKIEEYKKEITMDKNQKVLPAVAASKGLLAHPSNSVYVQFLSQLGPFKEALSDD-----VT 237
DB 178 EKEIADVKKQABASAN-NALVVLVNEGQLSAYGKSRFGLIHDTEGFAA--DDKIEAST 234
QY 238 KGLSKYLGYPYLOMNTETLSQVNPMEFIM-TNKASSNEPSLKELEKDPVWKLANVKNQ 296
DB 235 HGQS-----VSYYVLEKNPGILFVVDRTKRAIGDDSKDNVAANELIQTDAKND 285
QY 297 RVDILDRDLWARSRLISSEEMAKELVELSK 328
DB 286 KVIMLPQDVWYLSGGGLESMLH---MIEDVKK 314

RESULT 11

US-09-134-000C-4342
; Sequence 4342, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778

```
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4342
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4342

Query Match      16.7%; Score 283; DB 4; Length 327;
Best Local Similarity 30.4%; Pred. No. 2.1e-15;
Matches 101; Conservative 59; Mismatches 130; Indels 42; Gaps 13;

Qy 11 LSVIGLLFVLATACGNN-----SSNSSKESKDGVEIKHEEGTTKVPKPKRVVVLEYS 67
Db 17 VSMGLML-----SACQTNKKTADSATTETAKTEVTVDKTNQGLTVPKPKRVVFDNG 72
Qy 68 FVDALVALDV--KPGVIADNNKRNRIKPLRDKIGKTYTSVTRKOPNLEEISKLKPLII 125
Db 73 SLDTMDALGVDGVVGAFTKN-----IPAYLKTKYQVESAGGIKEPDLKINQLKPDII 127
Qy 126 ADNRHKGIVKDLNKAFTIEL-----KSPDGDYNNENIDAFKTIKALGKEEGKRLPEH 181
Db 128 I-SGRQDQYQQLKAIATPIYLAVDKPNWASTKQNIETLGTI---PDKKEVAKETIGL 183
Qy 182 DKKEEYKKEITMDKNQKVLPAVAAGSGLLAHPSNYSYVQFLSQLGFKALSD-----VT 237
Db 184 EKEIADVKKQAEASAN-NALVVLVNEGQLSAYGKGRFLIHDTFGFKAA--DDKIEAST 240
Qy 238 KGLSKYLKGPVLOMNTETLSQVNPFRMFIM-TNKASSNEPSLKELEKDPVWKKLNAVKNQ 296
Db 241 HQQS-----VSYEVLKNGPILFVVDRTKAIIGDDSKDNVAANELIKTDAGKND 291
Qy 297 RVDILDRDLWARSRGLISSEEMAKELVELSKK 328
Db 292 KVINLQPDVWYLSGGGLSEML---MIEDVKK 320

RESULT 12
US-09-543-681A-5565
; Sequence 5565, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5565
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5565

Query Match      16.5%; Score 278.5; DB 4; Length 325;
Best Local Similarity 28.0%; Pred. No. 4.9e-15;
Matches 94; Conservative 65; Mismatches 132; Indels 45; Gaps 11;

Qy 9 KILSVIGLLFVLIIATAACGNSS-----SNSKESKDGVEIKHEEGTTKVPKPKRVVVL 64
Db 12 KILAVTASLLSVSFSALSNTSVYQPNQIVEQQNDKVKVILKHLIGETAVSKNSKVLF 71
Qy 65 EYSFVDALVALDVK-----PUGIADNNKRNRIKPLRDKIGKTYTSVTRKOPNLEEISK 118
Db 72 DFLGVDLSVALGLADKVVGLPLGNAPYIKRIANDV-----ANVGKMAPFERLAE 124
Qy 119 LKPDIIADNRHKGIVKDLNKAFTIELKSPDGDYNNENIDA-FKTIKALGKEEGKRL 177
Db 125 LKPDIIII-TGRQGSYEKLTNIAPTINLTGNSANYLSDSVEANIQLIGELFDKQATQTQ 183

Qy 178 LEHDHKKIEBKKEITMDKNQKVLPAVAAGSGLLAHPSNYSYVQFLSQLGFKALSDVDT 237
Db 184 LATLNTVIEQAQK-AAGSDKKVLLVLLHNAAGNLN- -PNN-----QSVIYDVV 227
Qy 238 KGLSKYL-----KGPVLOMNTETLSQVNPFRMFIMTNKASSNEPSLKE--LEKDPVWK 288
Db 228 KAKKAEELPPVAEEKGRVVTSEMIKAVNPVDVFIIDRSEAIAGAGKLEKNTFENDEV-K 286
Qy 289 KLNKAVKNQVRDILDRDLWARSRGLISSEEMAKELVE 324
Db 287 TTSAYKGGIVVLQSLDWLWYLSGGGLSILTOQIEAVQ 322

RESULT 13
US-09-071-035-40
; Sequence 40, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-40

Query Match      15.9%; Score 268; DB 4; Length 296;
Best Local Similarity 30.3%; Pred. No. 3.2e-14;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

Qy 26 CGNN-----SSNSSKESKDGVEIKHEEGTTKVPKPKRVVVLEYSFVDALVALDV--KPV 80
Db 1 CQTNKKTADSATTETAKTEVTVDKTNQGLTVPKPKRVVFDNGSLDTMDALGVDGVV 60
Qy 81 GIADNNKRNRIKPLRDKIGKTYTSVTRKOPNLEEISKLKPLIIADNRHKGIVKDLNK 140
Db 61 GAPTKN-----IPAYLKTKYQVESAGGIKEPDLKINQLKPDIIII-SGRQDQYQQLKA 114
Qy 141 IAPTIEL-----KSPDGDYNNENIDAFKTIKALGKEEGKRLPEHKKIEBKKEITMDK 196
Db 115 IAPTIIYLAVDKPNWASTKQNIETLGTI---FDKEEVAKEKITGLEKEIADVKKQAEASA 171
```


;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 4860
;; LENGTH: 60
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4860

Query Match 15.7%; Score 265; DB 3; Length 60;
Best Local Similarity 92.9%; Pred. No. 6.3e-15;
Matches 52; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 280 BLEKDPVWKKLNAVKNQVRVDILDRDLWARSGLISSEEMAKELVELSKDKSKDKK 335
Db 5 QIRKRPVWKKLNAVKNQVRVDILDRDLWARSGLISSEEMAKELVELSKDKSKDKK 60

RESULT 16
US-09-583-110-3723
; Sequence 3723, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3723
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3723

Query Match 15.0%; Score 254; DB 4; Length 344;
Best Local Similarity 28.9%; Pred. No. 5.5e-13;
Matches 93; Conservative 58; Mismatches 115; Indels 56; Gaps 15;

Qy 13 VIGLLFVLIAACGNNS---SSNSKSSKSGDVEIKHEEGTTKVPKHPKRVVLEYSFV 69
Db 10 LIATLACIVFSACSSNSVKNENSTKSHAPDKIVLDHAFGQTILDKKPERVATIAGNH 69

Qy 70 DALVALDVKPGIADDN---KKNRIKP-----LRDKIGKYSVGTGRKQPNLEEISKLKP 121
Db 70 DVALALGIVPGFSKANYGVADKGVLPWTEKIKELNGKANLPDLDGLFAEALSNSKP 129

Qy 122 DLIADNRRHGI-----YKDLNKIAPTIELKSPDGYNNENI--DAFTISKALGKEEG- 174
Db 130 DVILAG---YSGITKEDYDTLSKIAPVAAVAKSPW---QTLWRDMIKIDSKALGMEKGD 183

Qy 175 -----KKRLEHDKTEEVKKEITMDKNQKVLPAV---AAKSGLLAHPNSYVGQ 221
Db 184 ELINKTEARISKLEKH-----PEIKGKI---KGKKVLPFTMIINAADTSKFVIYTKDPPAN 236

Qy 222 FLSQLG--FKBALSDDVTKGLSKYLKPGYLOMNTETLSQVNPFRMFMITNKASNEPSLK 279
Db 237 YLTDLGLVFPESLKEFESE--DSPAK-----EISAEANKINDADVII---TYGDDKTLE 286

Qy 280 BLEKDPVWKKLNAVKNQVRVDIL 301
Db 287 ALQKDPILGKINAIGNGAVAVI 308

RESULT 17
US-09-583-110-2973
; Sequence 2973, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2973
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2973

Query Match 14.8%; Score 250; DB 4; Length 321;
Best Local Similarity 27.2%; Pred. No. 1.1e-12;
Matches 92; Conservative 65; Mismatches 119; Indels 62; Gaps 14;

Qy 20 LIATAACGNNSSSNSKSSKSGDVE--IKHEEGTTKVPKHPKRVVLEYSFVDALVALDV 77
Db 17 LLLGACSTNSSTSQSTSSAPTEVTKSLDEVKLSKYPEKIVTDLGAADTIRALGF 76

Qy 78 KPGIADNKKRIKP-----LRDKIGKYSVGTGRKQPNLEEISKLKPDLLIADNRR 130
Db 77 -----EKNIVGMPTKTVTYKDLVGTKNVGFWMKEPDLRAIAALEPDLIIASPR 127

Qy 131 HKGIYKDLNKIAPTIELKSPDGYNNENIDA-FKTIISKALGK--BEKGKRLLEEHDKLEE 187
Db 128 QKFVDK-FKEIAPTIVLFOAKDDYWTSTKANIESLASAFGTGTQKAKEELAKLDKSIQE 186

Qy 188 Y-----KKEITMDKNQKVLPAVAAKSGLLAHPNSYVGQFLSQ-LGFKELSDDDVTK 238
Db 187 VATKNESDDKALAILNEGMAAFGAKSRP-----SFLYQTLKFKPT---DTTF 233

Qy 239 GLSKYLKPGYLOMNTETLSQVNPFRMFMITNK-----SSNEPSLKE--LEKDPVWKKL 290
Db 234 EDSRHGQ---EVSFESVKEINPDILFVINRTLAIGDSSNNNGVLENALIAETP----- 284

Qy 291 NAVKNQVRVDILDRDLWARSGLISSEEMAKELVELSKK 328
Db 285 -AAKNGKIIQLTPDLWYLSGGGLESTKLMIEDIOKALK 321

RESULT 18
US-09-107-433-4809
; Sequence 4809, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>

QY 220 -----GQFLSQ-LGFKEALSDDVTGKLSKYLKGPYLOMNTTTLQVNPFRMFWTN 269
Db 203 FMSVNRSGTLYQDLGLQ-----VPKVEEISKNATADQWVSLEKLAELADHIFVNS 259
QY 270 KASSNEPSLEKLEKDPVWKKNLNAVKNQVRDILDRDLWARSRLGISSEEMAKELVELSKK 328
Db 260 DESA-----PLFQEAIWKNLPAVKNNQVHTYDK-----KSSWLYNGPIANTQIVEDVKK 308

RESULT 22
US-08-961-083-24
; Sequence 24, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-24

Query Match 14.0%; Score 236; DB 3; Length 298;
Best Local Similarity 26.6%; Pred. No. 1.4e-11;
Matches 88; Conservative 65; Mismatches 112; Indels 66; Gaps 14;
QY 29 NSSNSSKSSKDGVE--IKHEGTTKVPKPKRVVVLVYSFVDALVALDVKPVGIADDN 86
Db 3 NSSTSQTSSAPTETVTKSSLDVSKVPEKIVTFDLGAADTIRALGF----- 53
QY 87 KQRIIKP-----LRDKIGKTVSGTRKQPNLEISKLKPDLLIADNNRHKGIYKDLN 139
Db 54 EKNVGMPTKTVPTLYKDLVGTGVKNGSMKEPDLAIAALEPDLIIASPRTKQFVDK-PK 112
QY 140 KIAPTIELKDFGDYNNIDA-FKTSKALGK--BEEGKRLREHDKKIEEY----- 188
Db 113 EIAPTVLFQASKDDYWTSTKANTIESLASAFGETGTQKAKEELTKLDSIQEVATKNESD 172
QY 189 KKEITMDKNQKVLPAVAKSGLLAHPNSYVGFQFLSQ-LGFK--EALSDDVTGKLSKYLK 245
Db 173 KKAALAILNEGKMAAFGAKSRF-----SFLYQTLKFKPTDTKFDSDRHG----- 216

QY 246 GPYLOMNTTTLQVNPFRMFWTNKA-----SNFPSLKE--LEKDPVWKKNLNAVKNQVR 297
Db 217 ---GVSFESVKEINPDILFVINRTLAIGGDSNDGVLENALIAETP-----AAKNGK 267
QY 298 VDILDRDLWARSRLGISSEEMAKELVELSKK 328
Db 268 IIQTPDLWYLSGGGLESTKLMIEDIQKALK 298
RESULT 23
US-09-536-784-24
; Sequence 24, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
; US-09-536-784-24

Query Match 14.0%; Score 236; DB 4; Length 298;
Best Local Similarity 26.6%; Pred. No. 1.4e-11;
Matches 88; Conservative 65; Mismatches 112; Indels 66; Gaps 14;
QY 29 NSSNSSKSSKDGVE--IKHEGTTKVPKPKRVVVLVYSFVDALVALDVKPVGIADDN 86
Db 3 NSSTSQTSSAPTETVTKSSLDVSKVPEKIVTFDLGAADTIRALGF----- 53
QY 87 KQRIIKP-----LRDKIGKTVSGTRKQPNLEISKLKPDLLIADNNRHKGIYKDLN 139
Db 54 EKNVGMPTKTVPTLYKDLVGTGVKNGSMKEPDLAIAALEPDLIIASPRTKQFVDK-PK 112
QY 140 KIAPTIELKDFGDYNNIDA-FKTSKALGK--BEEGKRLREHDKKIEEY----- 188
Db 113 EIAPTVLFQASKDDYWTSTKANTIESLASAFGETGTQKAKEELTKLDSIQEVATKNESD 172
QY 189 KKEITMDKNQKVLPAVAKSGLLAHPNSYVGFQFLSQ-LGFK--EALSDDVTGKLSKYLK 245
Db 173 KKAALAILNEGKMAAFGAKSRF-----SFLYQTLKFKPTDTKFDSDRHG----- 216

Qy	246	GPYLOMKTETLSQVNPERMFTWTKA-----SSNEPSLKE--LEKDPVWKLVNAVKQR	297
Db	217	---QEVSFESVKEINPDILFVINRTLATGGNSNDGVLENALIAETP-----AAKNGK	267
Qy	298	VDIIDLRLWARSRGLISSEEMAKELVELSKK	328
Db	268	IIOLTPLDLYLSSGGLESTKLMIEDIOKALK	298

```

RESULT 24
US-09-134-000C-6014
; Sequence 6014, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6014
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6014

```

RESULT 25
US-09-071-035-72
Sequence 72, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

```

Qy 246 GPYLQMTETLSQVNPERMFTWTKA-----SSNEPSLKE--LEKDPVWKLNNAVQKOR 299
Db 217 ---QEVSFESVKBINPDILFINRTLAIGDGNSSNDGVLENALIAETP-----AAKNGK 267
Qy 298 VDILDRDLWARSRGLISSEEMAKELVELSKK 328
Db 268 IIQTPLDWYLSGGLESTKLMIEDIQALK 298

RESULT 24
US-09-134-000C-6014
; Sequence 6014, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6014
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6014

Query Match 14.0%; Score 236; DB 4; Length 301;
Best Local Similarity 27.0%; Pred. No. 1.4e-11;
Matches 92; Conservative 65; Mismatches 94; Indels 90; Gaps 19

Qy 24 AACGN-NSSNSSKXSSKDGVEIKHEBGT-----KVPKHKPRVVVLEYSFV-DALVALD 76
Db 11 AACGNTNKEANNADKT-----HEVDTTLGNKVTVPAPKRIIA---SYLEDYLVALG 59

Qy 77 VKPGV---IADDNKKRIIKPLDKIGKYSVGTGKQNL-----EETSKLKPDLIIAD 127
Db 60 EKPVAQMTVGGGSIQDYLAKEKLD-----VPTISVDLPYEAVLKPEPDLLIS 107

Qy 128 NNR---HKGIYKDLANKIAPTIELKSPDGYNENI---DAFKTISKALGKEEKGKRLLEHD 182
Db 108 SSALVEGGKYKEYSKIATYVVYXN-----GENVTWRDQLEDIATVLDKKEQAKVLEDDYD 162

Qy 183 ---KKIEY--KKETITMDKNOKVLPAAVAKSGLLAHPNSYV-----GQFTSQ-IG 227
Db 163 TLTGKVQBYLKKD-----AGKSAVLWVTNNQVFMVSDNRSSGTVLYQDLG 209

Qy 228 PKEALSDDDVTGLSKYLKGPYLQMMTETLSQVNPERMFTWTKASSNEBPSLKELEKDPVW 287
Db 210 LQ---VPKLVEISIKNTADNQNQVLEKLAELDADHIFLVNSDESA-----PLFQEAIW 260

Qy 288 KKLNAVKNQRVILDRDLWARSRGLISSEEMAKELVELSKK 328
Db 261 KNI.PAVKNNQVHTYDK-----KSSWLYNGPIANTOIVEDVCK 297

```

RESULT 25
US-09-071-035-72
Sequence 72, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

RESULT 25
 US-09-071-035-72
 Sequence 72, Application US/09071035
 Patent No. 6448043
 GENERAL INFORMATION:
 APPLICANT: Gil H. Choi
 TITLE OF INVENTION: Enterococcus faecalis polynucleotides and Polypeptides
 NUMBER OF SEQUENCES: 496
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA

PRIOR FILING DATE: 1999-01-14
NUMBER OF SEQ ID NOS: 195
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 321
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-09-302-626B-12

Query Match 12.8%; Score 217; DB 4; Length 321;
Best Local Similarity 26.2%; Pred. No. 5.7e-10;
Matches 85; Conservative 63; Mismatches 136; Indels 40; Gaps 14;

QY 21 IATAAGC--NNSSNSSKSSKD-----GVBEIKHEEGTKVPKPKRNVVLEYSFVDA 71
DB 13 LALGACSPQNSDPAQAEQVSAQAQSEGVSVTVKTARGDVQIPQNPRIAYVDLGLMDT 72

QY 72 LVALDVKPGVGIADDDKNRIIKPLRDKIGKITS-VGTRKQPNLEISKLKPLDIIADNRR 130
DB 73 LSKLGVK-TGLSVD--KNRL--PYLEEFKTKPAGTLPEDPYETLNAYKPOLIIIGSRA 127

QY 131 HKGIYKOLNKIAPTIELKSFDDYNNEN-----IDAPKTIKALGKEEGKRLLEHDKKIE 186
DB 128 AKAPDK-LNEIAPTIENTADTANLKESAKERIDA---LAQIFGKQAEADKLKAEIDASPE 183

QY 187 EYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVGFSLQFKEALSDDTVKGLSKYLKG 246
DB 184 A-AKTAAGQKGLVILVNGGKMSAFSPSSRLGGWLHKDYGVPV--DEAIK-----EG 234

QY 247 PYLQ-MNTETLSQVNPFRMIFMTNKASSNE--PSLKELEKDPVWKKLNAVKNQVVDILDR 303
DB 235 SHGQPSIFPYLKEKNPDLFVLDRAAIGEGQAQKDVLDNPLVAETTTAKKGGQVYVLP 294

QY 304 DLWASRGLISSEEMAKELVELSKK 328
DB 295 ETYLAAGG-----AQELLNASKQ 312

RESULT 27
US-09-302-626B-10
Sequence 10, Application US/09302626B
Patent No. 6709660
GENERAL INFORMATION:
APPLICANT: Scalato, Enzo
APPLICANT: Masignani, Vega
APPLICANT: Rappuoli, Rino
APPLICANT: Pizza, Mariagrazia
APPLICANT: Grandi, Guido
TITLE OF INVENTION: Meningococcal Antigens
FILE REFERENCE: CHIR0159
CURRENT APPLICATION NUMBER: US/09/302,626B
CURRENT FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: PCT/IB99/00103
PRIOR FILING DATE: 1999-01-14
NUMBER OF SEQ ID NOS: 195
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 321
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-09-302-626B-10

Query Match 12.8%; Score 217; DB 4; Length 321;
Best Local Similarity 26.2%; Pred. No. 5.7e-10;
Matches 85; Conservative 63; Mismatches 136; Indels 40; Gaps 14;

QY 21 IATAAGC--NNSSNSSKSSKD-----GVBEIKHEEGTKVPKPKRNVVLEYSFVDA 71
DB 13 LALGACSPQNSDPAQAEQVSAQAQSEGVSVTVKTARGDVQIPQNPRIAYVDLGLMDT 72

QY 72 LVALDVKPGVGIADDDKNRIIKPLRDKIGKITS-VGTRKQPNLEISKLKPLDIIADNRR 130
DB 73 LSKLGVK-TGLSVD--KNRL--PYLEEFKTKPAGTLPEDPYETLNAYKPOLIIIGSRA 127

QY 131 HKGIYKOLNKIAPTIELKSFDDYNNEN-----IDAPKTIKALGKEEGKRLLEHDKKIE 186
DB 128 AKAPDK-LNEIAPTIENTADTANLKESAKERIDA---LAQIFGKQAEADKLKAEIDASPE 183

QY 187 EYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVGFSLQFKEALSDDTVKGLSKYLKG 246
DB 184 A-AKTAAGQKGLVILVNGGKMSAFSPSSRLGGWLHKDYGVPV--DEAIK-----EG 234

QY 247 PYLQ-MNTETLSQVNPFRMIFMTNKASSNE--PSLKELEKDPVWKKLNAVKNQVVDILDR 303
DB 235 SHGQPSIFPYLKEKNPDLFVLDRAAIGEGQAQKDVLDNPLVAETTTAKKGGQVYVLP 294

QY 304 DLWASRGLISSEEMAKELVELSKK 328
DB 295 ETYLAAGG-----AQELLNASKQ 312

RESULT 28
US-09-107-433-4519
Sequence 4519, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4519:
SEQUENCE CHARACTERISTICS:
LENGTH: 295 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...295
SEQUENCE DESCRIPTION: SEQ ID NO: 4519:
US-09-107-433-4519

Query Match 12.4%; Score 209; DB 4; Length 295;
Best Local Similarity 28.0%; Pred. No. 2.3e-09;
Matches 85; Conservative 53; Mismatches 110; Indels 56; Gaps 15;

Qy	13	VIGLLFVLIIATAACNNNS---SSNSKSSXGDGVEIKHEEGTKVYKQPKRVVVULEYSFV	69
Db	10	LIATLACIVFSACSSNSVQNEBTSKSHAPDKVLIDHAFQGTILDKKPVRVATTIAWGNH	69
Qy	70	DALVALOVKPGVIADDN---KKNRILKP-----LRDKTGYTSVGTRKQPNLEEISKLKP	121
Db	70	DVALALGIVPGFSKANYGVSAADKGVLPTWTEBKIKELNGKANLFPDLDGLNFEAISNSKP	129
Qy	122	DLIITADNNRHGKI---YKDLNKIAPTTELKSFDCGDYNEI---DAFKTISKALGKEEGB-	174
Db	130	DVILIAG---YSGITKEDYDTLSKJATPAVAAYSKPM---QTLMWRDMIKIDSKALGMEKGB	183
Qy	175	-----KKRLTEHDKTIEYKKEITMDKNQKVLPAV---AAKSGLLAHPSNSYVQG	221
Db	184	ELIKNTTEARISKELEKH---PEIKGKI---KGKVLFTMINAATDSKFWIYTSKDXPRAN	236
Qy	222	FLSQLG--FKREALSDDVTKGLSKYLGPLYQMNTTLSQVNPFRMFIMTKNASSNEPSJK	279
Db	237	YLTDLGLVPFSPSLKEFBSE--DSPAK---EISAEAEANKINDADVII----TVGDDKTLE	286
Qy	280	ELEK 283	
Db	287	ALOK 290	

```

RESULT 29
US-09-710-279-1888
Sequence 1888, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1888
LENGTH: 299
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-09-710-279-1888

```

```

QY      292 AVKNQVRVDILDRD 304
      |||| |:::|
Db      264 AVXNTRIETVDGD 276

RESULT 30
US-09-134-001C-5667
; Sequence 5667, Application US/09134001C
; Patent NO. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5667
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5667

```

RESULT 31
US-09-107-532A-5819
; Sequence 5819, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Walcham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
;


```

Db      73  LSKLGVK-TGLSDV--KNRL--PYLEEYFKTKPAGTLFEPDDETLYNAYKPOLIIIGRA 127
Qy      131 HKGIYKDLNKIAPTIELKSFQGDYNN-----IDAFKTIKSGALGKEEKGKRLRLEHDKKIE 186
Db      128 AKAFDK-LNEIAPTIENTADTANLKESAKERIDA---LAQIFGKKAEADKLKAEIDASFSE 183
Qy      187 EYK 189
      |
Db      184 AAK 186

RESULT 34
US-09-489-039A-13095
Sequence 13095, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13095
LENGTH: 337
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13095

```

```

; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/IB99/00103
; PRIOR FILING DATE: 1999-01-14
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (142)..(143)
; OTHER INFORMATION: unknown
;
US-09-302-626B-8

Query Match          9.1%; Score 153; DB 4; Length 164;
Best Local Similarity 30.1%; Pred. No. 4.3e-05;
Matches 47; Conservative 31; Mismatches 62; Indels 16; Gaps 7;

Qy 21 IATAACG--NNSSNSGSKESKD-----GVEIKHEEGTTKVPKPKRVVVLSEYSFVDA 71
Db 13 LALGACSPQNSAPQAQKAVSAQAQTEGASVTVKTARGDVQIQPQNERIAVYDLGMLDT 72
Qy 72 LVALDVKPGVIADDNKNRIIKPLRDKIGKVT--VGTRKQPNLEEISKLKPDLLIADNRR 130
Db 73 LSLKGVK-TGLSVD--KNRL--PYLEEYFKTKTPAGTLFEPDYETINAYKPOLIIGSRA 127
Qy 131 HKGIYKDLNKIAPTIELKSPDGDYNNENIDAPKTIK 166
Db 128 AKAFDK-LNEIAPTIXXTATNLKESAKEASTILAQ 162

RESULT 36
US-09-489-039A-7331
; Sequence 7331, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7331
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
;
US-09-489-039A-7331

Query Match          9.0%; Score 152; DB 4; Length 397;
Best Local Similarity 22.3%; Pred. No. 0.00017;
Matches 69; Conservative 61; Mismatches 119; Indels 60; Gaps 13;

Qy 54 VPKHPRVWVLEYSFVDALVALDVKP--VGIADDNKNRIIKP---LRDKIGKYTSVGT 107
Db 83 LPQHTRVVVLQHTQLLLVQLHAAEDIVGVLSWSQKQ--LGQFARPMPEIGQLATPGD 140
Qy 108 RKQPNLEEISKLKPDL-----IIADNNRHGIYKDLNKIAPTII- 145
Db 141 LTQVNTIESLLARPPQVVFVANYAPPAMIAIQQAGIPVALSLRHDA--GEKNQNPMTMA 199
Qy 146 -ELKSPDGDYNNENIDAPKTIKSKALGKEEGKKRIEHDKKIEEYKKEIT-----MDKNQK 199
Db 200 DEEQVYNAGLMEGI-----RLTGEVVERQPEAEALHTYTFEARKQANAFVADIPPNQR 252
Qy 200 VLPVAAKSGLLAHPNSVYQFLSQLGFKREALSDVTKGLSKYLVKGPYLQMWTTETLSQV 259
Db 253 V-RVYMANPDLNTYGAKYITGLMMAHAGALNVAASV-KGAR-----QVSLQVLEW 302
Qy 260 NPERMFWTNTKASGNEPSLKELEKDPVWKLNNAVKNQRVDILDRDLWARSGLISSEEMA 319

```

Db 303 NPOVIFVQDRYPQ-----VVKQIENDPQWQIDAIVKXHRVWLMP-----YAKAWGYMPPEALA 356
QY 320 KELVELSKK 328
Db 357 LGELWMAKK 365

RESULT 37
US-09-489-039A-12270
; Sequence 12270, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12270
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12270

Query Match 8.5%; Score 144; DB 4; Length 353;
Best Local Similarity 20.9%; Pred. No. 0.00067;
Matches 67; Conservative 52; Mismatches 137; Indels 64; Gaps 10;
QY 26 CGNSSSSSSKSSKSGDGVKIEKHEEGTKVPKH-----57
Db 19 CPDAGGTDQRHS--GTDLRHPDGDPAIPGRGAGKCLLMNPTLITRRLLIAMALS 76
QY 58 -----PKRVVVLVYSFVDALVDALVDKVPVGIAD-DNKKRIIPK-LRDKIG 100
Db 77 PLLQMRGAQAADVDPQVRVVALEWLPAAELLALGVTPYGVADIPNVLWVNEPALPDSV- 135
QY 101 KYTSVTRKQPNLEISKLPDLIIADNNRHGIYKDLNKAIPATTELKSPDGDYENIDA 160
Db 136 --IDVGLRTEPNLELLTKMPSPFVMSAGYSPSEK-LARIAPIRGFTFSDGKRLPMAQ 192
QY 161 PRTISA--LCGEEGKKLEEDHKIEYKKEITMDKNQKVLPAVAKSGLLAHPNS 217
Db 193 RSLLENADLLGKTQQAQRHLAEFDALMESLRPRFAGRGDRPLMLTSLDPRHVLVFGNC 252
QY 218 YVQFSLQGLFKEALSDVTKGLSKYLKGPYLMQNTETLSQVNPFRMFMINTKASSNEPS 277
Db 253 LFOEVLDRFGIKNAWTHGEAAFWGS-----VSVGIDRLAAFNADVICFDH---GNERD 302
QY 278 LKELEKDPVWKKLVNAVKNQR 297
Db 303 MAQLLATPLWQAMPFVRAGR 322

RESULT 38
US-09-489-039A-7324
; Sequence 7324, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7324
; LENGTH: 369

; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7324
Query Match 8.5%; Score 143.5; DB 4; Length 369;
Best Local Similarity 21.9%; Pred. No. 0.00079;
Matches 56; Conservative 47; Mismatches 90; Indels 63; Gaps 11;
QY 51 TTKVPKHPKRVVVL-EYSFVDALVALDVKPV-----GIADNKK 88
Db 84 TTVVPLHPQIRVSMHDLDTITPLIELGAPPIASHGRTRPDGSHYLRSSAQLTGVDNDSD 143
QY 89 NRIIKPLRDKIGKTSVTRKQPNLEISKLPDLIIADNNRHGIYKDLNKAIPATTELK 148
Db 144 IRFI-----GT-ADIDLEAAARPDIIITEPSRHVSF-EQLEKIAPIATVSID 188
QY 149 SFDGDYENIDAFKTIKALKEEGKKRLEEDHKIEYKKEITMDKN--QKVLPAVA 206
Db 189 HLQSAPE---LYRKLALQLTGTQ-----PRLAILERRVQEQIKQLKAMVNPQYSVSVIOA 241
QY 207 KSG-LLAHPNSYVQFSLQGLFKEALSDVTKGLSKYLKGPYLMQNTETLSQV 260
Db 242 NNGKVTVHSHYHALGRVLRDAGFRPPLIERIPD-----GQRIDVSAEQLELPD 290
QY 261 PERFMINTKASSNEP 276
Db 291 ADFVFATWRSDTGGKP 306

RESULT 39
US-09-328-352-7262
; Sequence 7262, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7262
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7262

Query Match 8.3%; Score 140.5; DB 4; Length 354;
Best Local Similarity 22.9%; Pred. No. 0.0013;
Matches 75; Conservative 56; Mismatches 113; Indels 83; Gaps 15;
QY 10 ILSVIGLLFVLIAAACGNNSSSSKSSKSGDGVKIEKHEEGTK--VPKHPKRVVVLVYS 67
Db 15 IMKLASLTLILLCGSSSEONRKTQENSTSHSCVFDSTNTKVCVAKPAQRIYSLFES 74
QY 68 FVDALVALDV--KPVGIADD-----NKKNRRIKPLRDKIGKTSVTRKQPNLEIS 117
Db 75 GLDGLYMLGQGDKVIGIPAEVVIQPLLEFNAYSKIDKRIANKQLAAPSQGA-NATNIESIV 133
QY 118 KLKPLDIIADNNRHGIYKDLNKAIPATTE-LKSF-----DGDYENIDAFKTIKAL 168
Db 134 LLKPLDLIVVSGQ-----TQTIELLQFGIAVYVMSGTYKQVKESLSIALIS 182
QY 169 GKEEGKKRLEEDHKIEYKKEITMDKNQKVLPAVAKSGLLAHPNSYV---GQFLS 224
Db 183 GAQKAQAILNFSDEIVAE-----VAAKTARQPNKQSIYYAWSGRIFS 226
QY 225 QLGKFEALSDVTKGLSKYLKGPY-----LQMTTETLSQVNPFRMFMINTKASSNE 275
Db 227 TSG-RESITNDFIE-----LAGAYIVQTNANQPNVNPETLIEWPNDIVLM-----NTN 275
QY 276 PSL-----KELEKDPVWKKLVNAVKNQR 298

Db 276 PKLIYERKELQ-----GLSAVQNRKV 296

RESULT 40
 US-08-961-083-194
 ; Sequence 194, Application US/08961083
 ; Patent No. 6159469
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et. al.
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
 ; NUMBER OF SEQUENCES: 452
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961.083
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 194:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 132 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-961-083-194

Query Match 7.9%; Score 134; DB 3; Length 132;
 Best Local Similarity 30.5%; Pred No. 0.0012;
 Matches 39; Conservative 23; Mismatches 52; Indels 14; Gaps 4;
 QY 21 IATAACGNNS---SSNSKSSKDGVEIKHEGTTKPKPKRVVVLVYSFVDALVALDV 77
 Db 2 IVFSACSSNSKNEENTSKEHAPDKIVLDHAFGQTILDKPervATIAWGNHHDVALALGI 61
 QY 78 KPVGIADN---KQRIKP-----LRDKIGKYSVGTGRKQPNLEEISKLKPDLLIADNN 129
 Db 62 VPVGFSKANYGVSDKGVLPWTEEKIKSLNGKANLFDLGLNFEAISNSKPDVILAG-- 119
 QY 130 RHKGIYKD 137
 Db 120 -YSGITKE 126

Search completed: August 25, 2005, 09:27:38
 Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:17:46 ; Search time 65 seconds
(without alignments)

2639.177 Million cell updates/sec

Title: US-10-724-972A-6352

Perfect score: 1690

Sequence: 1 GVESVRLKILSVIGLLFVL.....EEMAKVELSKDSKDNK 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1681	99.5	334	2	Q8CNG3	Q8CNG3 staphylococ
2	1232	72.9	327	2	Q99S89	Q99S89 staphylococ
3	1232	72.9	327	2	Q7A499	Q7A499 staphylococ
4	1230	72.8	327	2	Q6GQ2	Q6GQ2 staphylococ
5	1227	72.6	327	2	Q8NVD6	Q8NVD6 staphylococ
6	1227	72.6	327	2	Q6G7D7	Q6G7D7 staphylococ
7	725.5	42.9	319	2	Q8NKA6	Q8NKA6 staphylococ
8	725.5	42.9	319	2	Q6GAH5	Q6GAH5 staphylococ
9	724.5	42.9	319	2	Q99V55	Q99V55 staphylococ
10	724.5	42.9	319	2	Q7A6B6	Q7A6B6 staphylococ
11	719.5	42.6	319	2	Q6G145	Q6G145 staphylococ
12	679.5	40.2	315	2	Q34348	Q34348 bacillus su
13	611	36.2	348	2	O07616	O07616 bacillus su
14	608.5	36.0	325	2	Q651P5	Q651P5 bacillus li
15	516.5	30.6	298	2	Q9CBB7	Q9CBB7 pasteurella
16	488.5	28.9	307	2	Q8GRB3	Q8GRB3 vibrio para
17	487.5	28.8	307	2	Q87FM4	Q87FM4 vibrio para
18	476	28.2	302	2	Q7M254	Q7M254 photorhabdu
19	476	28.2	304	2	Q6D898	Q6D898 erwinia car
20	463	27.4	300	2	Q931F9	Q931F9 shigella fl
21	460	27.2	300	1	FECS_ECOLI	FECS_ECOLI
22	454.5	26.9	306	2	Q88916	Q88916 pseudomonas
23	415.5	24.6	325	2	Q65F62	Q65F62 bacillus li
24	405	24.0	330	2	Q87491	Q87491 staphylococ
25	405	24.0	330	2	Q7A127	Q7A127 staphylococ
26	405	24.0	330	2	Q7A2Y7	Q7A2Y7 staphylococ
27	405	24.0	330	2	Q7A869	Q7A869 staphylococ
28	405	24.0	330	2	Q6GD10	Q6GD10 staphylococ
29	405	24.0	330	2	Q6GKJ0	Q6GKJ0 staphylococ
30	404.5	23.9	321	2	Q81V85	Q81V85 bacillus an
31	404	23.9	321	2	Q6HJ99	Q6HJ99 bacillus th

32 399 23.6 321 2 Q63G31 Q63G31 bacillus ce
33 397 23.5 321 2 Q73DM9 Q73DM9 bacillus ce
34 396.5 23.5 322 2 Q81I09 Q81I09 bacillus ce
35 392 23.2 324 2 Q817L3 Q817L3 bacillus ce
36 391 23.1 320 2 Q81DY8 Q81DY8 bacillus ce
37 391 23.1 324 2 Q81L65 Q81L65 bacillus th
38 389 23.0 324 2 Q6HCZ7 Q6HCZ7 bacillus li
39 380 22.5 315 2 Q65NE2 Q65NE2 bacillus li
40 379.5 22.5 315 2 Q65ENG Q65ENG bacillus li
41 379.5 22.5 321 2 Q738W0 Q738W0 bacillus ce
42 378 22.4 316 2 Q8XM98 Q8XM98 clostridium
43 378 22.4 320 2 Q81R05 Q81R05 bacillus an
44 377 22.3 320 2 Q6HJA9 Q6HJA9 bacillus th
45 373 22.1 320 2 Q63BT9 Q63BT9 bacillus ce

ALIGNMENTS

RESULT 1

Q8CNG3 PRELIMINARY; PRT; 334 AA.
AC Q8CNG3;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Ferrichrome ABC transporter.
CN OrderedLocName=SEL768;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
DR EMBL; AE016750; AA05409.1; -
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2_1.
KW Complete proteome.
SQ SEQUENCE 334 AA; 37438 MW; 92C02D41A5ABAF33 CRC64;

Query Match 99.5%; Score 1681; DB 2; Length 334;
Best Local Similarity 99.7%; Pred. No. 5.2e-84;
Matches 333; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 VESVRLKILSVIGLLFVLIIATAACNNSSNSSKESKDGVEIKHEEGTTKVPKPKRV 61
Db 1 MVESVRLKILSVIGLLFVLIIATAACNNSSNSSKESKDGVEIKHEEGTTKVPKPKRV 60
QY 62 VVLESFVDALVALDVKPGIADNNKNRIIKPLRDKIGKTVTSVGRKQPNLEETSKLP 121
Db 61 VVLESFVDALVALDVKPGIADNNKNRIIKPLRDKIGKTVTSVGRKQPNLEETSKLP 120
QY 122 DLIIADNNRHGIYKDLNKIAPTIELKSPDGYNNENIDAFKTIISKALGKEEGKKLEEH 181
Db 121 DLIIADNNRHGIYKDLNKIAPTIELKSPDGYNNENIDAFKTIISKALGKEEGKKLEEH 180
QY 182 DKKIEEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYVQFISQLGFKALSDDVTGKLS 241
Db 181 DKKIEEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYVQFISQLGFKALSDDVTGKLS 240
QY 242 KYLGKPYLQMTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKLVNAVKQNRVDIL 301
Db 241 KYLGKPYLQMTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKLVNAVKQNRVDIL 300

```

Qy 302 DRDLWARSRLGISSEEMAKELVELSKKDKXK 335
Db 301 DRDLWARSRLGISSEEMAKELVELSKKDKXK 334

RESULT 2
Q99889 PRELIMINARY; PRT; 327 AA.
ID Q99889
AC Q99889;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to ferrichrome ABC transporter.
GN OrderedLocusNames=SAV2177;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
RL Lancet 357:1225-1240(2001).
RX EMBL; AP003364; BAB58339.1; -.
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Ouchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T., Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
RL Lancet 357:1225-1240(2001).
RX EMBL; AP003364; BAB58339.1; -.
DR PIR; D90013; D90013.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 327 AA; 36591 MW; 6AECEBC4E17CDB7 CRC64;

Query Match / 72.9%; Score 1232; DB 2; Length 327;
Best Local Similarity 71.9%; Pred. No. 1.5e-59;
Matches 235; Conservative 42; Mismatches 50; Indels 0; Gaps 0;

Qy 5 VRLKILSVIGLLFVLVIAAAGNNSNSSKESKOGVEIKHBEETTKVPKPKRVVVL 64
Db 1 MRGLKTFSLGLIALLVAAAGNTDSSKESSTKDTISVKDENGTVKVPKDAKRIIVL 60

Qy 65 BYSPFDALVALDVKPGVGIADNKKRIIKPLDKIKGYTSVGTTRKQPNLEISKLKPLI 124
Db 61 EYSPADALAALDVKPGVGIADNKKRIIKPLDKIKGYTSVGTTRKQPNLEISKLKPLI 120

Qy 125 IADNRHKGIVKDLNKIAPTIELSKFGDYNENIDAPKTIKSKALGKEEGKRLLEHDKK 184
Db 121 IADSRHKGINKELNKIAPTLSLKSFGDYKQNTINSFKTIKALNKEKEGKRLAEHDKL 180

Qy 185 IEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYVQGFLSQGFKEALSDDVTGKLSKYL 244
Db 181 INKYKDEIKPDRNQKVLPAVAAKAGLLAHPNSYVQGFLNELGFKNALSDDDVTGKLSKYL 240

Qy 245 KGPYQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPWKKLNVAQNQRVDILDRD 304
Db 241 KGPYQLDLTEHLADLNPERMIIMTDHAKKDSAEFKKLQEDATWKKLNVAQNQRVDILDRD 300

Qy 305 LWARSRLGISSEEMAKELVELSKKDKSK 331
Db 301 VWARSRLGISSEEMAKELVELSKKQK 327

RESULT 3
Q7A499 PRELIMINARY; PRT; 327 AA.
ID Q7A499
AC Q7A499;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to ferrichrome ABC transporter.
GN OrderedLocusNames=SAV2177;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
RL Lancet 357:1225-1240(2001).
RX EMBL; AP003364; BAB58339.1; -.
DR PIR; D90013; D90013.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 327 AA; 36591 MW; 6AECEBC4E17CDB7 CRC64;

Query Match / 72.9%; Score 1232; DB 2; Length 327;
Best Local Similarity 71.9%; Pred. No. 1.5e-59;
Matches 235; Conservative 42; Mismatches 50; Indels 0; Gaps 0;

Qy 5 VRLKILSVIGLLFVLVIAAAGNNSNSSKESKOGVEIKHBEETTKVPKPKRVVVL 64
Db 1 MRGLKTFSLGLIALLVAAAGNTDSSKESSTKDTISVKDENGTVKVPKDAKRIIVL 60

Qy 65 BYSPFDALVALDVKPGVGIADNKKRIIKPLDKIKGYTSVGTTRKQPNLEISKLKPLI 124
Db 61 EYSPADALAALDVKPGVGIADNKKRIIKPLDKIKGYTSVGTTRKQPNLEISKLKPLI 120

Qy 125 IADNRHKGIVKDLNKIAPTIELSKFGDYNENIDAPKTIKSKALGKEEGKRLLEHDKK 184
Db 121 IADSRHKGINKELNKIAPTLSLKSFGDYKQNTINSFKTIKALNKEKEGKRLAEHDKL 180

Qy 185 IEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYVQGFLSQGFKEALSDDVTGKLSKYL 244
Db 181 INKYKDEIKPDRNQKVLPAVAAKAGLLAHPNSYVQGFLNELGFKNALSDDDVTGKLSKYL 240

Qy 245 KGPYQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPWKKLNVAQNQRVDILDRD 304
Db 241 KGPYQLDLTEHLADLNPERMIIMTDHAKKDSAEFKKLQEDATWKKLNVAQNQRVDILDRD 300

Qy 305 LWARSRLGISSEEMAKELVELSKKDKSK 331
Db 301 VWARSRLGISSEEMAKELVELSKKQK 327

RESULT 4
Q6GE02 PRELIMINARY; PRT; 327 AA.
ID Q6GE02
AC Q6GE02;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative transport system binding lipoprotein.
GN OrderedLocusNames=SAR2268;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;

```

RA Holden M.T.G., Feil E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
 RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
 RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
 RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
 RA Spratt B.G., Parkhill J.,
 RT "Complete genomes of two clinical *Staphylococcus aureus* strains:
 RT evidence for the rapid evolution of virulence and drug resistance.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 DR EMBL; BX571856; CAG41246.1; -.
 DR GO; GO:0005381; F:iron ion transporter activity; IEA.
 DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
 DR InterPro; IPR002491; Peripla_BP.
 DR Pfam; PF01497; Peripla_BP_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 327 AA; 36687 MW; AC34D0F769D71265 CRC64;

Query Match 72.8%; Score 1230; DB 2; Length 327;
 Best Local Similarity 72.2%; Pred. No. 1.9e-59;
 Matches 236; Conservative 39; Mismatches 52; Indels 0; Gaps 0;

QY 5 VRGLKILSVIGLLFVLIATAACGNSSSSSSKSSKSGVEIKHEEGTTKVPKPKRVVVL 64
 DB 1 MRGLKTFSLGLIALLVAACGNTDSSKESSTKDTISVDENGTVKVPKDAKRIIVL 60

QY 65 EYSFVDALVALDVKPVGIAADNKKRIIKPLRDKIKGYTSVGTTRQPNLEISKLKPDLI 124
 DB 61 EYSFADALVALDVKPVGIAADNKKRIIKPVREKIDYTSVGTTRQPNLEISKLKPDLI 120

QY 125 IADNRHKGIIYKDLNKIAPTIELKSPDGYNNENIDAFKTIKALGKEEGKRLAEHDKK 184
 DB 121 IADSRHKGINKELNKIAPTIELKSPDGYKQNIIDAFKTIKALDKKEGKRLAEHDKL 180

QY 185 IEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQFLSGLFKEALSDDVTGKLSKYL 244
 DB 181 INKYKDEIFDRNQKVLPAVAKAGLLAHPNSYVQFLNELGFKNALSDDVTGKLSKYL 240

QY 245 KGPYLQMTETLSQVNPFRMIFMTNKASNPESLKELEKDPVWKLNNAVQNORVDILDRD 304
 DB 241 KGPYLQDTEHLADLNPERMIIMTDNAKDSAEFKLQEDATWKKLNNAVKNRVDIVDRD 300

QY 305 LWASRGLISSEEMAKELVELSKDSK 331
 DB 301 VWASRGLISSEEMAKELVELSKKEQK 327

RESULT 5
 Q8NV06 PRELIMINARY; PRT; 327 AA.
 AC Q8NV06;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE M2103 protein.
 GN OrderedLocusNames=MW2103;
 OS *Staphylococcus aureus* (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.
 OC NCBI_TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MW2;
 RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
 RX Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA";
 RL Lancet 359:1819-1827(2002).
 DR EMBL; AP004829; BAB95968.1; -.
 DR GO; GO:0005381; F:iron ion transporter activity; IEA.

DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
 DR InterPro; IPR002491; Peripla_BP.
 DR Pfam; PF01497; Peripla_BP_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 327 AA; 36577 MW; 6AEDCDBC4E169CB7 CRC64;

Query Match 72.6%; Score 1227; DB 2; Length 327;
 Best Local Similarity 71.6%; Pred. No. 2.8e-59;
 Matches 234; Conservative 42; Mismatches 51; Indels 0; Gaps 0;

QY 5 VRGLKILSVIGLLFVLIATAACGNSSSSSSKSSKSGVEIKHEEGTTKVPKPKRVVVL 64
 DB 1 MRGLKTFSLGLIALLVAACGNTDSSKESSTKDTISVDENGTVKVPKDAKRIIVL 60

QY 65 EYSFVDALVALDVKPVGIAADNKKRIIKPLRDKIKGYTSVGTTRQPNLEISKLKPDLI 124
 DB 61 EYSFADALVALDVKPVGIAADNKKRIIKPVREKIDYTSVGTTRQPNLEISKLKPDLI 120

QY 125 IADNRHKGIIYKDLNKIAPTIELKSPDGYNNENIDAFKTIKALGKEEGKRLAEHDKK 184
 DB 121 IADSRHKGINKELNKIAPTIELKSPDGYKQNIIDAFKTIKALNKKEGKRLAEHDKL 180

QY 185 IEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQFLSGLFKEALSDDVTGKLSKYL 244
 DB 181 INKYKDEIFDRNQKVLPAVAKAGLLAHPNSYVQFLNELGFKNALSDDVTGKLSKYL 240

QY 245 KGPYLQMTETLSQVNPFRMIFMTNKASNPESLKELEKDPVWKLNNAVQNORVDILDRD 304
 DB 241 KGPYLQDTEHLADLNPERMIIMTDNAKDSAEFKLQEDATWKKLNNAVKNRVDIVDRD 300

QY 305 LWASRGLISSEEMAKELVELSKDSK 331
 DB 301 VWASRGLISSEEMAKELVELSKKEQK 327

RESULT 6
 Q6G7D7 PRELIMINARY; PRT; 327 AA.
 AC Q6G7D7;
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Putative transport system binding lipoprotein.
 GN OrderedLocusNames=SA52078;
 OS *Staphylococcus aureus* (strain MSSA476).
 OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.
 OC NCBI_TaxID=282459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
 RX Holden M.T.G., Feil E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
 RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
 RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
 RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
 RA Spratt B.G., Parkhill J.;
 RT "Complete genomes of two clinical *Staphylococcus aureus* strains:
 RT evidence for the rapid evolution of virulence and drug resistance.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 DR EMBL; BX571857; CAG43886.1; -.
 DR GO; GO:0005381; F:iron ion transporter activity; IEA.
 DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
 DR InterPro; IPR002491; Peripla_BP.
 DR Pfam; PF01497; Peripla_BP_2; 1.
 KW Complete proteome; Lipoprotein.
 SQ SEQUENCE 327 AA; 36577 MW; 6AEDCDBC4E169CB7 CRC64;

Query Match 72.6%; Score 1227; DB 2; Length 327;
 Best Local Similarity 71.6%; Pred. No. 2.8e-59;
 Matches 234; Conservative 42; Mismatches 51; Indels 0; Gaps 0;

```
Qy 5 VRGKILSVIGLLFVLITATAAGNNSNSSKSSKDGVEIKHEGTTKVPKPKRVVVL 64
Db 1 MRGLKTSFSLGLVALLVAACGNTDSSKSSSTKDTISVKDENGTVKVPKDAKRIIVL 60
Qy 65 EYSPFDALVALDVKPVGIADGNKNRIIKPDRDKIGKTSVGTQKPNLEISIKLPDLI 124
Db 61 EYSPADALVALDVKPVGIADGNKKRIIKPREKIGDYTSVGTQKPNLEISIKLPDLI 120
Qy 125 IADNRHKGIIYKDLNKIAPTIELKSPDGYNYENIDAFKTSKALGKEBEGKRLLEHDKK 184
Db 121 IADSSRHKGINKELNKIAPTIELKSPDGYDKONINSFKTIAKALNKEKEGKRLAEHDKL 180
Qy 185 IEEYKKEITMDKNQKVLPAVAAGSGLLAHPNSYVGQFLSQLGKEALSDDVTGLSKYL 244
Db 181 INKYKDEIKFDRNQKVLPAVAKAGLLAHPNSYVGQFLNELGFKNALSSDDVTGLSKYL 240
Qy 245 KGPVLQMTETLSOVNPERMFTMTNKASSNPSLKELEKDPVWKLANAVKNQVRDILDRD 304
Db 241 KGPVLQDTEHLADINPERMFTMTDHAQKDSAEFKLQSDATWKKLANAVKNRVDIVDRD 300
Qy 305 LWASRGLISSEEMAKELVELSKDSK 331
Db 301 VWASRGLISSEEMAKELVELSKKEQK 327

RESULT 7
Q8NXXA6
ID Q8NXXA6 PRELIMINARY; PRT; 319 AA.
AC Q8NXXA6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MW0921 protein.
GN OrderedLocusNames=MM0921;
OS Staphylococcus aureus (strain MW2);
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004825; BAB94786.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR InterPro; IPR002491; PeriPla_BP.
DR Pfam; PF01497; PeriPla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 319 AA; 35552 MW; DB386C3BB2C60535 CRC64;

Query Match 42.9%; Score 725.5; DB 2; Length 319;
Best Local Similarity 46.9%; Pred. No. 5.7e-32;
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

Qy 13 VIGLLFVLITATAAGNNSNSSKSSKDGVEIKHEGTTKVPKPKRVVVLVYSFVDAL 72
Db 9 VVFMILVAVAGCGQKDT-----EKTMTTIKDELGTETKIKKNPKRVVVLVYSFADYL 63
Qy 73 VALDVKPVGIADGNKNRIIKPLRDKIGKTSVGTQKPNLEISIKLPDLIADNRHK 132
Db 64 AALDMKPVGIADGSSKNITKSVRDKIGAYESVGRPOPNEVISEKLPDLIADVSRHK 123
Qy 133 GIYKDLNKIAPTIELKSPDGYNYENIDAFKTSKALGKEBEGKRLLEHDKKIEYKKEI 192
Db 124 KIKSLSKIAPTITMLVSGTGDYNANIETAFKTVAKAVGKEGKEKLEKHKDLAEIRKI 183
Qy 193 TMDKNQKVLPAVAAGSGLLAHPNSYVGQFLSQLGKEALSDDVTGLSKYLKGPYLQWN 252
Db 184 EQSTLKSAPAFGISRAGFMFINNETFMGQFLKMGIOPEVTKDTHVGERKGGPYIYN 243
Qy 253 TETLSOVNPERMFTMTN-KASSNPSLKELEKDP-VWKLANAVKNQVRDILDRDLWASR 310
Db 244 NEELANINPKVMILATDGTDKNRTKF-----IDPAVWKS LKAVKDKNKYVDVDRNKLKSR 299
Qy 311 GLISSEEMAKELVELSKK 328
Db 300 GTIASMESMAEDLEKIAEK 317
```

```
Db 184 EQSTLKSAPAFGISRAGFMFINNETFMGQFLKMGIOPEVTKDTHVGERKGGPYIYN 243
Qy 253 TETLSOVNPERMFTMTN-KASSNPSLKELEKDP-VWKLANAVKNQVRDILDRDLWASR 310
Db 244 NEELANINPKVMILATDGTDKNRTKF-----IDPAVWKS LKAVKDKNKYVDVDRNKLKSR 299
Qy 311 GLISSEEMAKELVELSKK 328
Db 300 GTIASMESMAEDLEKIAEK 317

RESULT 8
Q6GAH5
ID Q6GAH5 PRELIMINARY; PRT; 319 AA.
AC Q6GAH5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transport system extracellular binding lipoprotein.
GN OrderedLocusNames=SAS0973;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=1521324; DOI=10.1073/pnas.0402521101;
RA Hoiden M.T.G., Feil E.J., Lindsay J.A., Pescock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571857; CAG42748.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR InterPro; IPR002491; PeriPla_BP.
DR Pfam; PF01497; PeriPla_BP_2; 1.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 319 AA; 35552 MW; DB386C3BB2C60535 CRC64;

Query Match 42.9%; Score 725.5; DB 2; Length 319;
Best Local Similarity 46.9%; Pred. No. 5.7e-32;
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

Qy 13 VIGLLFVLITATAAGNNSNSSKSSKDGVEIKHEGTTKVPKPKRVVVLVYSFVDAL 72
Db 9 VVFMILVAVAGCGQKDT-----EKTMTTIKDELGTETKIKKNPKRVVVLVYSFADYL 63
Qy 73 VALDVKPVGIADGNKNRIIKPLRDKIGKTSVGTQKPNLEISIKLPDLIADNRHK 132
Db 64 AALDMKPVGIADGSSKNITKSVRDKIGAYESVGRPOPNEVISEKLPDLIADVSRHK 123
Qy 133 GIYKDLNKIAPTIELKSPDGYNYENIDAFKTSKALGKEBEGKRLLEHDKKIEYKKEI 192
Db 124 KIKSLSKIAPTITMLVSGTGDYNANIETAFKTVAKAVGKEGKEKLEKHKDLAEIRKI 183
Qy 193 TMDKNQKVLPAVAAGSGLLAHPNSYVGQFLSQLGKEALSDDVTGLSKYLKGPYLQWN 252
Db 184 EQSTLKSAPAFGISRAGFMFINNETFMGQFLKMGIOPEVTKDTHVGERKGGPYIYN 243
Qy 253 TETLSOVNPERMFTMTN-KASSNPSLKELEKDP-VWKLANAVKNQVRDILDRDLWASR 310
Db 244 NEELANINPKVMILATDGTDKNRTKF-----IDPAVWKS LKAVKDKNKYVDVDRNKLKSR 299
Qy 311 GLISSEEMAKELVELSKK 328
Db 300 GTIASMESMAEDLEKIAEK 317
```

DB 300 GIIASESMAEDLEKIAEK 317

RESULT 9

Q99V55 ID Q99V55 PRELIMINARY; PRT; 319 AA.

AC Q99V55;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 26, Last annotation update)

DE Similar to ferrichrome ABC transporter.

GN OrderedLocusNames=SAV1038;

OS Staphylococcus aureus (strain Mu50 / ATCC 700699).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158878;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Mu50 / ATCC 700699;

RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T., Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.

RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."

RL Lancet 357:1225-1240(2001).

DR EMBL; AP003361; BAB57200.1; -.

DR PIR; E89872; E89872.

DR GO; GO:0005381; P:iron ion transporter activity; IEA.

DR GO; GO:0006827; P:high affinity iron ion transport; IEA.

DR InterPro; IPR002491; Peripla_BP.

DR Pfam; PF01497; Peripla_BP_2; 1.

KW Complete proteome.

SQ SEQUENCE 319 AA; 35566 MW; DE6D7C7AB4831574 CRC64;

Query Match 42.9%; Score 724.5; DB 2; Length 319;

Best Local Similarity 46.9%; Pred. No. 6.5e-32;

Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

QY 13 VIGLLFVLIIATAACGNSSSSSKSSKSGDVEIKHEEGTTKVPKHPKRVVVLVLEYSFVDAL 72

DB 9 VVFMILLVAVAGCGQKDT-----EKEETMTIKDELGTEKIKKPKRVVVLVLEYSFADYL 63

QY 73 VALDVPVGIADNDKKNRIKPLRDKIGKITSVGTGKQPNLEIEISKLPDLIIADNNRHK 132

DB 64 AALDMKPVGIADGSGTKNITKSVDRKIGAYESVGSRPQNMNEVISKLPDLIIADVSRHK 123

QY 133 GIYKDLNKIAPTIELKSFQDYNENIDAFKTIKSKALGKEEGKRLLEHDKKIEEYKKEI 192

DB 124 KIKSELSKIAPTIMLVSGTGDYNANIEAFKTVAKAVGKEKEGKLEKHKILAEIRKKI 183

QY 193 TMDKNQKVLPAVAAKSGLLAHPNSYVGQFLSQLGFKALSDDVTGKLSKYLKGPYLOWN 252

DB 184 EQSTLSKSAFAGISRAGFINNEDTFMGQFLIKMGIQPEVTKDKTTHVGERKGGPYIYN 243

QY 253 TETLSQVNPFRMFMNTN-KASSNEPSLKELEKDP-VWKKLNNAVKNQVRVILDRDLWARSR 310

DB 244 NEELANINPKVMILATDGKTDKNRTKF----IDPAVWKSLSKAVKDNKVYDVRNKLKSR 299

QY 311 GLISSEEMAKELVELSKK 328

DB 300 GIIASESMAEDLEKIAEK 317

RESULT 11

Q6GI45 ID Q6GI45 PRELIMINARY; PRT; 319 AA.

AC Q6GI45;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Transport system extracellular binding lipoprotein.

GN OrderedLocusNames=SAR1011;

OS Staphylococcus aureus (strain MRS252).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=282458;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=15213324; DOI=10.1073/pnas.0402521101;

RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J., Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,

DB 300 GIIASESMAEDLEKIAEK 317

RESULT 10

Q7A6B6 ID Q7A6B6 PRELIMINARY; PRT; 319 AA.

AC Q7A6B6;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

```

RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hannan Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.,
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance."
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG40015.1; -.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 319 AA; 35566 MW; 5885A83E3F502ABE CRC64;

Query Match 42.6%; Score 719.5; DB 2; Length 319;
Best Local Similarity 46.2%; Pred. No. 1.2e-31;
Matches 147; Conservative 65; Mismatches 95; Indels 11; Gaps 4;

QY 13 VIGLFLVLIATAACGNNSSNSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVSFDAL 72
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 9 VVFMILVVAVAGCQKDT-----EKTMTTIKDELCTEKIKKPKRIVVLVSFDYL 63
QY 73 VALDVKPGIADDNKKNIIPKLRDKIGKYSVGTGRKQPNLEISKLPDLIIADNNRHK 132
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 AALDKMPGVIADGSGTKNITKSVRDKIGAYESVGRSQPNMNEVSKLPDLIIADVSRHK 123
QY 133 GIYDKLNKIAPTIELKSPGDYNNENIDAPKTSIKALGKEEGKKELEHDKKIEYKKEI 192
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 KIKSELSKIAPTILVSGTDYNNANIEAPTKAVAKVGKEGKLEKHNIKILAIRKKI 183
QY 193 TMDKNQKVLPAVAAGSLLAHPNSNVGQFLSQLGFEALSDDVTGKLSKYLKGPYLOVN 252
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 EQSTLSAPAFAGISAGFINNEDTFMGQFLIKMGIQPEVTKDHTAHRGKGGPIYLN 243
QY 253 TETLSQVNPFRIMTN-KASSNEPSLEKEDP-VWKKLNKAVNQKRVDDILDRDLWASR 310
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 NEELANIPKVMILATNGTKDKNRTKFT---IDPAVMKSLKAVKONKVVVDNRKWLQSR 299

QY 311 GLISSEENAKVELSKK 328
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 GIMASESMAEDLEKIAEK 317

RESULT 12
Q34348 PRELIMINARY; PRT; 315 AA.
ID O34348 Q79ET3;
AC O34348; Q79ET3;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE YfmC protein.
GN Name=yfmc; OrderedLocusNames=BSU07520;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]_TaxID=1423;
RP SEQUENCE FROM N.A.
RC STRAIN=166;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Gutseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

```

```

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Konigstein G., Krohn S.,
RA Kunano M., Kurita K., Lapides A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maue C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone P., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Takakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.P., Zumstein E.,
RA Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AC327;
RX MEDLINE=97417488; PubMed=9272861; DOI=10.1016/S0378-1119(97)00130-3;
RA Yamamoto H., Uchiyama S., Nugroho F.A., Sekiguchi J.,
RT "Cloning and sequencing of a 35.7 kb in the 70 degree-73 degree region
RT of the Bacillus subtilis genome reveal genes for a new two-component
RT system, three spore germination proteins, an iron uptake system and a
RT general stress response protein."
RL Gene 194:191-199(1997).
DR EMBL; Z99108; CAB12561.1; -.
DR EMBL; D86417; BAA22317.1; -.
DR PIR; B69812; B69812.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 315 AA; 35020 MW; 2DA3BFFC49EFC39 CRC64;

Query Match 40.2%; Score 679.5; DB 2; Length 315;
Best Local Similarity 47.0%; Pred. No. 1.8e-29;
Matches 149; Conservative 50; Mismatches 103; Indels 15; Gaps 4;

QY 8 LKILSVIGLFLVLIATAACGNNSSNSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVS 67
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 9 IAIMSVL-LLACLIVSGCSSQNNNGSGKSDSRVHDEEGKTVSGTKRVVVLVS 67
QY 68 FVDALVALDVKPGIADDNKKNIIPKLRDKIGKYSVGTGRKQPNLEISKLPDLIIAD 127
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 FLDAVHNLGITPVGIADDNKKMIKLVGSSI-DYTSVGTSEPNLEVISLKPDLIIAD 126
QY 128 NNRHKGIVKLNKIAPTIELKSPGDYNNENIDAPKTSIKALGKEEGKKELEHDKKIEE 187
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 AERHKNYIKQLKKIAPTIELKSRATYDETIDTSFTIAKLNKEDGKLEAHEKVKVIND 186
QY 188 YKKEITMDKNQKVLPAVAAGSLLAHPNSNVGQFLSQLGFEALSDDVTGKLSKYLKGP 247
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 LKAEIPKDNENIVLVARADSFPQLHTSSSYDGEIFKMLGFTTHAVKSD-----NA 236
QY 248 YLQWNTYLSQVNPFRIMTN-KASSNEPSLEKEDP-VWKKLNKAVNQKRVDDILDRDLWA 307
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 YQEVSELSKIDPDLIFISANEKGT----IVDEWKNTPLNKLNKAVKNGQVYDADRDTWT 293
QY 308 RSRGLISSEENAKELIVE 324
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 RFRGKISSETSAKDVLK 310

RESULT 13
O07616

```


DR InterPro; IPR002491; Peripla_BP.
KW Pfam; PF01497; Peripla_BP_2; 1.
KX Complete proteome.
SQ SEQUENCE 307 AA; 34185 MW; DDA125C81E0E1638 CRC64;

Query Match 28.8%; Score 487.5; DB 2; Length 307;
Best Local Similarity 35.4%; Pred. No. 5.1e-19;
Matches 110; Conservative 72; Mismatches 106; Indels 23; Gaps 5;

QY 16 LFLVLIATACGNNSSNSSKSSKDGVIKHEEGTTKVPHKPVVVLEYSFVDALVAL 75
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
15 VLLMLSSALMSFSASQAR-----SVQDEGTTFELEAIPQRIVLVESFVDALA 66
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 76 DVKPVGIADNKKRIIPLRKDKIGKYTVSGTRKQNPLEBIEISKLKPDIIADN 135
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 67 DVSPVGADDNDAIRVIPAVRAKIEPWQSGMRSQPSLEIAVLKPDIIADAERHCA 126
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 136 KDANKTAPIELKSPGDYNENIDAPKTISKALGKEEGKLEEDHKIEYKKEITMD 195
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 127 QDLQRIAPTLLLSRGTEYOENLSAQKIGVAIGKQAQMTOIRIELHKOTMAEFKH 184
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 196 KNOKVLPAVAASKGLLAHPNSVYGQFLSOLGFKEALSDDVTKGLSKYLKGYLQ 255
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 185 TQTIOFGVSDKGWHLSPVSAGGVLTSLGIOSLPASE-----RNAYITSFEL 236
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 256 LSQVNPERPIMTNKASSNEPSL-KBLEKDPVWKLNNAVKNQVRDILDRDLWAR 314
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 237 LLKNTDPDWLLV----GLYSQPNIVDSWRNRNPFLKLTASKKOQLVEVSPELS 292
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 315 SEEMAKELVEL 325
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 293 AEEIARNLEAL 303

RESULT 18
Q7MZ54 PRELIMINARY; PRT; 302 AA.

ID Q7MZ54 AC Q7MZ54 DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Iron(III) dicitrate-binding periplasmic protein FecB.
GN Names=fecB; OrderedLocusNames=plu4447;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Ruaniok C., Frangeul L., Buchrieser C., Givaudan A.-F., Taoutit S., Boes S., Bouraux-Eude C., Chandler M., Charles J.-F., Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S., Medigue C., Lanois A., Powell K., Sigurdson P., Vincent R., Wingate V., Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RA "The genome sequence of the entomopathogenic bacterium Photorhabdus luminescens.";
RT Nat. Biotechnol. 21:1307-1313(2003).
RL EMBL; BX571873; CAE1681.1; -;
DR Photolyst; plu4447; -;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 302 AA; 33546 MW; 579D5013C7627F8F CRC64;

Query Match 28.2%; Score 476; DB 2; Length 302;
Best Local Similarity 36.0%; Pred. No. 2.1e-18;
Matches 102; Conservative 60; Mismatches 107; Indels 14; Gaps 3;

QY 43 VEINKHEEGTTKVPHKPVVVLEYSFVDALVALDVKPVGIADNKKRIIPLRKDKIG 102

```

Db 200 SVLAALGL-----SVPKPIN---NAPMASINLEQLAINPQWLIVTHYR---EBSIVKR 247
Qy 281 LEKDPVKKLVAVKQNRVDILDRDLWARSRLGISSEEMAKELVELSK 327
Db 248 WQDDTLWNLEBAQQKQQAADVSDNWARMRGIFAAERVGSDDTKKIF 294

RESULT 20
ID Q93F19 PRELIMINARY; PRT; 300 AA.
AC Q93F19,
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FeCB.
GN Name=FeCB;
OS Shigella flexneri 2a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=42897;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21437601; PubMed=11553538;
RX DOI=10.1128/JAI.69.10.6012-6021.2001;
RA Luck S.N., Turner S.A., Rajakumar K., Sakellariis H., Adler B.;
RT "Ferric dicitrate transport system (Fec) of Shigella flexneri 2a
RT YSH6000 is encoded on a novel pathogenicity island carrying multiple
RT antibiotic resistance genes.";
RL Infect. Immun. 69:6012-6021(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Luck S.N., Turner S.A., Rajakumar K.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF326777; AAL08454.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
SQ SEQUENCE 300 AA; 33166 MW; 6E0B0CF5B99A4AAD CRC64;

Query Match 27.4%; Score 463; DB 2; Length 300;
Best Local Similarity 35.5%; Pred. No. 1.1e-17;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;

Qy 45 IKHEEGTTKPKPKRVVVLVLEYSFVDALVDPVGIADNKKNRILKPLRDKIGKYTS 104
Db 24 VQDEHGTFTLEKTPQRIWVLSLSPADALAAVDVSPIGIADNDKAKRILPEVRAHLKPQOS 83

Qy 105 VQTRKQPNLEISKLKPLDIADNRRHKGIVKDLNKAIPTELKSFDDGYNENIDAFKTI 164
Db 84 VQTRAQPSLEAIAALKPDLIIADSSRHAGVYALQQIAPVLLKSRNETHYAEVLQSAAI 143

Qy 165 SKALGKEEGKRLKEHDKITBEYKKEITMDKNQKLVPAVAKSGLLAHPNSVYGVQFLS 224
Db 144 GEMVKKKEEMQARLEQHKERNAQWASQ--LPKGTVAFTSREQFNLTQETWGSVLA 201

Qy 225 QLGFKEALSDVDVKGSLKYLKPYL-QMNTETLSQVNPFRPMTNKASSNEPSPSLKLEK 283
Db 202 SLGL-----NVPAAMAGSSMPSIGLEQLAVNPALWLAHYR---EBSIVKRWQ 248

Qy 284 DPVKKLVAVKQNRVDILDRDLWARSRLGISSEEMAKELVEL 325
Db 249 DPLWQMLTAAQKQQAASVDSNTWARMRGIFAAERIAADTVKI 290

RESULT 21
FECB_ECOLI
ID FECB_ECOLI STANDARD; PRT; 300 AA.
AC P15028; P76816;
DT 01-APR-1990 (Rel. 14, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Iron(III) dicitrate-binding periplasmic protein precursor.

```

```

GN Name=fecB; OrderedLocusNames=b4290;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacter;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=89213950; PubMed=2651410;
RA Staudenmaier H., van Hove B., Yaraghi Z., Braun V.;
RT "Nucleotide sequences of the fecBCDE genes and local
RT proteins suggest a periplasmic-binding-protein-dependent
RT mechanism for iron(III) dicitrate in Escherichia coli.";
RL J. Bacteriol. 171:2626-2633(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
CC -!- FUNCTION: Binds citrate-dependent iron(III); part of the binding-
CC protein-dependent transport system for uptake of citrate-dependent
CC iron(III).
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- SIMILARITY: Belongs to the bacterial extracellular solute-binding
CC protein family 8.
CC -!- SIMILARITY: Contains 1 iron siderophore/cobalamin periplasmic-
CC binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M26397; AAA23762.1; -.
CC EMBL; U14003; AAA97186.1; ALT INIT.
CC EMBL; U00096; AAC77246.1; ALT_INIT.
CC PIR; S56515; QRECD3.
CC EchoBase; EB0283; -.
CC EcoGene; EG10287; fecB.
CC InterPro; IPR002491; Peripla_BP.
CC Pfam; PF01497; Peripla_BP_2; 1.
CC PROSITE; PS00983; FE_B2_FBP; 1.
KW Complete proteome; Iron transport; Periplasmic; Signal; Transport.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 300 Iron(III) dicitrate-binding periplasmic
FT protein.
FT DOMAIN 39 295 Fe/B12 periplasmic-binding.
FT VARIANT 8 8 L -> V (in strain B).
FT CONFLICT 23 23 T -> M (in Ref. 1).
FT CONFLICT 57 57 I -> S (in Ref. 1).
SQ SEQUENCE 300 AA; 33146 MW; 6024ED0F9C82D0EA CRC64;

Query Match 27.2%; Score 460; DB 1; Length 300;
Best Local Similarity 35.5%; Pred. No. 1.6e-17;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;

Qy 45 IKHEEGTTKPKPKRVVVLVLEYSFVDALVDPVGIADNKKNRILKPLRDKIGKYTS 104
Db 24 VQDEHGTFTLEKTPQRIWVLSLSPADALAAVDVSPIGIADNDKAKRILPEVRAHLKPQOS 83

Qy 105 VQTRKQPNLEISKLKPLDIADNRRHKGIVKDLNKAIPTELKSFDDGYNENIDAFKTI 164
Db 84 VQTRAQPSLEAIAALKPDLIIADSSRHAGVYALQQIAPVLLKSRNETHYAEVLQSAAI 143

Qy 165 SKALGKEEGKRLKEHDKITBEYKKEITMDKNQKLVPAVAKSGLLAHPNSVYGVQFLS 224

```

Db 144 GEMVGGKREMOARLEQHKERMAQWASQ--LPKGTTRVAFGTSREQQFNHLHTQETWTGSLVA 201
 QY 225 QLGPKREALSDVTKGLSKYLGKPYL-QMNTETLSQVNPFRMFTMTKASSNEPSLKELEK 283
 Db 202 SLGL-----NVPAMAGASPSIGLEQLLAVNPFWLLVAHYR---EESIVRRWQ 248
 QY 284 DPVWKKLNAVKNQVRDILDRDLWARSRLGSLISEEMAKELVEL 325
 Db 249 DPLWQMLTAAQKQVAVSDNTWARMRGIFAAERIAADTVKI 290

RESULT 22
 Q88916 PRELIMINARY; PRT; 306 AA.
 AC Q88916;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Iron(III) dicitrate transport system, periplasmic iron-binding protein
 DE PecB.
 GN Name=fecB; OrderedLocusNames=PSPT00763;
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
 RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
 RA Madupu R., Dougherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
 RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
 RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
 RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
 RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
 RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
 RA Bender C.L., White O., Fraser C.M., Collier A.;
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen
 RT Pseudomonas syringae pv. tomato DC3000.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
 DR EMBL; AB016858; AA054305.1; -;
 DR TIGR; PSPT00763; -;
 DR GO; GO:0005381; F:iron ion transporter activity; IEA.
 DR GO; GO:0005827; P:high affinity iron ion transport; IEA.
 DR InterPro; IPR002491; Peripla_BP.
 DR Pfam; PF01497; Peripla_BP_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 306 AA; 32872 MW; 371BA711BC7A6108 CRC64;

Query Match 26.9%; Score 454.5; DB 2; Length 306;
 Best Local Similarity 34.6%; Pred. No. 3.2e-17;
 Matches 112; Conservative 60; Mismatches 111; Indels 41; Gaps 7;
 QY 11 LSVIGLLFVLIAATAACNNSSNSSKSSKSGDVGVEIKHEEGTKV--PKHPKRVVLEYSF 68
 Db 9 LLACGLL-TLLSAA-----QAAPIDIDGQHKVLPDTPKRVVLEYSF 52
 QY 69 VDVALVDKPVGIADNDKNNRIIPKRDKI GYTSVGRKQPNLEEISKLKPDLLIADN 128
 Db 53 LDGLASVGVTPVGAADGDASRVLPKRVKAVGEMQSVGLRSQPNIEVIARLKPDLIIADL 112
 QY 129 NRHKGIVKDLNKIAPTIELKSPDGDYNNIDAFKTSKALGKEEGKKELEHDKKIEY 188
 Db 113 GRHQALYNDLASLAPTLMLPSRGEDYQGLSKSAGLIGMALGKGPENQARIAENRHLKTV 172
 QY 189 KKEITWDKQKVLPAVAAGSGLLAHPSPSVGQFSLQGFKEALSDDVTGKLSKYLKGPY 248
 Db 173 AEQIPADSN--VLFVGAREDSFVHGPHSYAGSVLQAIG-----LQVPE 214
 QY 249 LQMTETLSQVNPFRMFTMTK-----ASSNEPSLKELEKDPVWKKLNAVKNQVRDILDR 303
 Db 215 VRNNAAPTEFVSLQQLALDPNLLVGHVRRPSIVDTWSKQPLWQVLGAVRNQKQVAEVDG 274

QY 304 DLWARSRLGSLISEEMAKELVELSK 327
 Db 275 DSWARNRGIMASEQIADDAALAVLK 298

RESULT 23
 Q65F62 PRELIMINARY; PRT; 325 AA.
 AC Q65F62;
 DT 25-OCT-2004 (TRENBLrel. 28, Created)
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE yfiY (Periplasmic binding protein).
 GN Name=yfiY; ORFNames=BL02129, BL103475;
 OS Bacillus licheniformis DSM 13.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=279010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 13;
 RX PubMed=15383718;
 RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baumeister S., Henne A., Liebegang H., Merkl R.,
 RA Ehrenreich A., Gottschalk G.;
 RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
 RT Organism with Great Industrial Potential.";
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14580;
 RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
 RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
 RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
 RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
 RA Berka R.M.;
 RT "Complete genome sequence of the industrial bacterium Bacillus
 RT licheniformis and comparisons with closely related Bacillus species.";
 RL Genome Biol. 5:R77-R77(2004).
 DR EMBL; AE017333; AAU42302.1; -;
 DR EMBL; CP000002; AAU4933.1; -;
 SQ SEQUENCE 325 AA; 36700 MW; 949283E3C6C24775 CRC64;

Query Match 24.6%; Score 415.5; DB 2; Length 325;
 Best Local Similarity 33.2%; Pred. No. 4.6e-15;
 Matches 108; Conservative 63; Mismatches 117; Indels 37; Gaps 10;
 QY 8 LKLSVIGLLFVLIAAT--AACNNSSNSSKSSKSGDVGVEIKHEEGT-TKVPKHPKRVVVL 64
 Db 1 MKRWSIVGFIALLAISILAAACGGKDDSAKGGKNDTITVKHAMGTEDQVPADPKRVVVL 60
 QY 65 EYSFVDALVALVDKPVGIADNDKNNRIIPKRDKI GYTSVGRKQPNLEEISKLKPDLLI 124
 Db 61 TNEGTEALLEIGIKPVGAVKSWGTGDPWYKHQIADMKGVKQIGLEGEPNIEAELKPDLLI 120
 QY 125 IADNNRHGIVKDLNKIAPTIELKSPDGDYNNIDAFKTSKALGKEEGKKELEHDKK 184
 Db 121 IGKRRHEKIYDKLEKIAPTTFSEITLGEWKDN---FSLYAKAVNKEEGKKELEHDKK 177
 QY 185 IEYKKEITMDKNQK--VLPVAAGSGLLAHPSPSVGQFSLQGFK-----EAL---SD 234
 Db 178 VSELSDKLGDQKKRVSVIRFVAGQSRI--YYKDSFPFGIILDLQGLFKYPEKQKLEFEKQD 235
 QY 235 DVTGKLSKYLKGPYLOMNTETLSQVNPFRMFTMTKASSNEPSLKELE-----KDPVW 287
 Db 236 D-----KPVFTENKESIPDMDGVLFYFTYKPAE---SKKEATQNDWTSDDLW 283
 QY 288 KKLNAVKNQVRDILDRDLWARSRL 312
 Db 284 KKLNAVKTGNAHEVDADATWTTAGGV 308

RESULT 24

```

087491
ID O87491 PRELIMINARY; PRT; 330 AA.
AC O87491;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Lipoprotein SirA.
GN Name=sirA;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ISP3;
RC MEDLINE=99175440; PubMed=10049373;
RA Heinrichs J.H., Gatlin L.E., Kunach C., Choi G.H., Hanson M.S.;
RT "Identification and characterization of SirA, an iron-regulated
protein from Staphylococcus aureus.";
RL J. Bacteriol. 181:1436-1443(1999).
[2]
RN SEQUENCE FROM N.A.
RP PubMed=14688077; DOI=10.1128/TAI.72.1.29-37.2004;
RA Dale S.E., Doherty-Kirby A., Lajoie G., Heinrichs D.E.;
RT "Role of siderophore biosynthesis in virulence of Staphylococcus
aureus: identification and characterization of genes involved in
production of a siderophore.";
RL Infect. Immun. 72:29-37(2004).
DR EMBL, AF079518; AAC62496.1; -.
DR EMBL, AY251022; AAP82062.1; -.
DR PIR, G89771; G89771.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006877; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_Bp.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Lipoprotein.
SQ SEQUENCE 330 AA; 36744 MW; 63710BA22B208F61 CRC64;

Query Match 24.0%; Score 405; DB 2; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.7e-14;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12

Qy 8 LKILSVIGLLFVLITATAAGNNSNSKSGKGVETKHBEGTTKVPKPKRVRVVLVYS 67
Db 5 IKML-VVTLAFLLV-LAGSGNSNQSSDNKDQKETTISKHANGTTETKIGKPKRVVTLVQG 62
Qy 68 FVDALVALDVKPVGIADNNKNRIKPRDKIGKTVTSVGRKQPNLEISKLPDLIAD 127
Db 63 ATDVAVSLGVKPVGAVESWTKQKPEYIKNDLKDTKIVGQEPAPNLEISKLPDLIVAS 122
Qy 128 NNRHKGIVKIDNKTAPTIELKSGFDGYNENIDAPKTIISKALGKBEGKRLLEHDKKTEE 187
Db 123 KVRNEKVDQLSKKAPT-V---STDVTFKFK-DTTKLMGKALQKEAEADLLKKYDDKVA 178
Qy 188 YKKEITMDKNQKVLPAVAASKGLLAHPNS-----SYVGQFSLQGFKEALSDVVTKGL 240
Db 179 FQK---DAKAKYKDAWFLKASVNVNFRADHTRIVAGGYAGTEINDLGFKR-----NKDL 228
Qy 241 SKYLKG--PYLQMT-ETLSQVNPFRMIMTNKASSNEPSL-----KELEKDPVWKKLNA 292
Db 229 QKQVDNGKDIIQLTSKESIPLMNADHIFVVKSDPNKDAALVKCTESEWTSSEWKKNLDA 288
Qy 293 VKNQRV-DILDRDLWARSGLLISSEMAKELVELSKDKSKDNK 335
Db 289 VKNNQVSDDLDEITWNLAGGYKSLKGLDLDLYE---KLNIEKQSK 330

RESULT 25
Q7A127
ID Q7A127 PRELIMINARY; PRT; 330 AA.
AC Q7A127;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)

```

[illegible]


```

Qy 188 YKKEITMDKQKVLPAVAAKSGLLAHPSN-----SYVGQFLSQLGKEALSDDVTKGL 240
Db 179 FQK-----DAKAKYKDAWFLKASVNVFRADHTRIIVAGGYAGIILNDLGFKR-----NKDL 228

Qy 241 SKYLKG--PYLQMT-ETLSQVNPFRMIFMTNKASSNEPSL-----KELEKDPVWKQLNA 292
Db 229 QKQVNDGKDIIQLTSKESIPLMNADHIPVVKSDPNKDAALVKKTESWTSKWKNLDA 288

Qy 293 VKNQV--DILDRDLWARSRLISSEMAKELVELSKDKSKDNK 335
Db 289 VKNQVSDDLDEITWNLAGGYKSKLKLDDLYE--KLNIEKQSK 330

RESULT 29
Q6GKJ0 PRELIMINARY; PRT; 330 AA.
ID Q6GKJ0
AC Q6GKJ0
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Lipoprotein.
GN Name=sirA; OrderedLocusNames=sAR0118;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=1521324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG39144.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR Complete proteome; Lipoprotein.
SQ SEQUENCE 330 AA; 36717 MW; F815203503849D6D CRC64;

Query Match 24.0%; Score 405; DB 2; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.7e-14;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

Qy 8 LKILSVIGLLPVLITATAACGNSSSSSSKESKDGVEIKHEGTTKPKPKRVVVLVYS 67
Db 5 IKML-VVTLAFLV-LACSGNSNKQSDSKDKETTSIKHANGTTEIKGPKRVVTLVQG 62

Qy 68 FVDALVALDVPVGIADDNKNRIIKPLRDKIGKYSVGTGRKQPNLEISKLPDLIAT 127
Db 63 ATDVAVSLGVKPVGAVESWTKPKFEYIKNDLKDKTKIVGQBPAPNLEISKLPDLIVAS 122

Qy 128 NNRHKGIYKNDKNTAPTILSKSPDGDYNNENIDAFKTSKALGKEBEGKKRLSEHDKLEE 187
Db 123 KVRNEKVVQDLSKIAPTY---STDVTVPFKF-DTTKLMGKALGKEAEDLLKKYDDKVA 178

Qy 188 YKKEITMDKQKVLPAVAAKSGLLAHPSN-----SYVGQFLSQLGKEALSDDVTKGL 240
Db 179 FQK-----DAKAKYKDAWFLKASVNVFRADHTRIIVAGGYAGIILNDLGFKR-----NKDL 228

Qy 241 SKYLKG--PYLQMT-ETLSQVNPFRMIFMTNKASSNEPSL-----KELEKDPVWKQLNA 292
Db 229 QKQVNDGKDIIQLTSKESIPLMNADHIPVVKSDPNKDAALVKKTESWTSKWKNLDA 288

```

```

Qy 293 VKNQV--DILDRDLWARSRLISSEMAKELVELSKDKSKDNK 335
Db 289 VKNQVSDDLDEITWNLAGGYKSKLKLDDLYE--KLNIEKQSK 330

RESULT 30
Q81V85 PRELIMINARY; PRT; 321 AA.
ID Q81V85
AC Q81V85; Q613H2; Q6KX85;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN OrderedLocusNames=BA0615; BAS0581; GBAA0615;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Faulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rikstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madupu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.D., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomas B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Ames / isolate 0581;
RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Sterne;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017026; AAP24632.1; -.
DR EMBL; AE017334; AAT29716.1; -.
DR EMBL; AE017225; AAT52909.1; -.
DR TIGR; BA0615; -.
DR TIGR; GBAA0615; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR Complete proteome.
SQ SEQUENCE 321 AA; 36302 MW; BB17CFF136D893CC CRC64;

Query Match 23.9%; Score 404.5; DB 2; Length 321;
Best Local Similarity 31.7%; Pred. No. 1.8e-14;
Matches 103; Conservative 73; Mismatches 120; Indels 29; Gaps 8;

Qy 13 VIGLFLVLIATAACGNSSSSSSKESKDGVEIKHEGTTKPKPKRVVVLVYSFVDA 71
Db 13 VFSLAFLSL-USACGKSNKESKEDTKEMIPVEHAMGKTEVPANPKRVVILTNEGTEA 71

Qy 72 IVALDVPVGIADDNKNRIIKPLRDKIGKYSVGTGRKQPNLEISKLPDLIADNNRH 131
Db 72 LLELGVKPVGAVSWTGDPMYPHIKDKMKVKVVGDEGVNVETIASLKPLDIIGNKMRH 131

```



```

QY 132 KGIYKDLNKIAPTIELKSGFDGYNENIDAFKTIKALGKEEGKRLBEHDKIEEYKKE 191
DB 132 EKVYEQKAIAPTVPSETLGEWKDN---FKFYAKALNKEKGQKVLAAAYDKRMKDLKAK 188
QY 192 ITMDKNQKVLPAVAAKSGLLAHPNSYVQFLSQLGFK-----EALSDVTGKLSKYLKGP 247
DB 189 LGDKVQNEISWVRFPMPGVDVRIYHGTFTSGVILKELGFKRPGDQNKDPAERNVSK----- 243
QY 248 YLQMTETLSQVNPFRMFTMTNKASSNEPSLKELEK---DPVWKKLNAVKNQVLDLRL 303
DB 244 -----ERISAMGDGVLFYFTFD-KGNEKKGSELEKEYINDPLFKNLNAVKNQKAYKVDD 296
QY 304 DLWARSGLISS-----EEMAKELVE 324
DB 297 VIVNTAGGVMAANLLDDIEKRFVK 321

RESULT 31
Q6HNJ9 PRELIMINARY; PRT; 321 AA.
ID Q6HNJ9
AC Q6HNJ9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron(III) dicitrate ABC transporter, periplasmic protein.
GN Name=fecB; OrderedLocusNames=Bt9727_0525;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27."
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017355; AAF2409.1; -
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 321 AA; 36373 MW; 478614F471C1CF3 CRC64;

Query Match 23.9%; Score 404; DB 2; Length 321;
Best Local Similarity 31.4%; Pred. No. 1.9e-14;
Matches 101; Conservative 73; Mismatches 120; Indels 28; Gaps 7;

QY 16 LLFVLIIATACGNSSNSK-SKDGVEIKHEGTTKVPKPKRVVLEYSFVDALVA 74
DB 15 LLAFSLLSACGKNTKESKEDTKKEMIPVHAMGKTEVPANPKRVVILNEGTEALLE 74
QY 75 LDVPGVGIADNKNRIIPKLDKIGKVTSGTRKQPNLEISKLKPLIIADNNRHGI 134
DB 75 LGVPGVAGKSWTGDPTWPHIKDKMKVKVGDGQVNVETIASLKPDLIIQNKVHEKV 134
QY 135 YKDLNKIAPTIELKSGFDGYNENIDAFKTIKALGKEEGKRLBEHDKIEEYKKEITM 194
DB 135 YEQLKAIAPTVPSETLGEWKDN---FKFYAKALNKEKGQKVLAAAYDKRMKDLKAKLGD 191
QY 195 DKNQKVLPAVAAKSGLLAHPNSYVQFLSQLGFK-----EALSDVTGKLSKYLKGPYLO 250
DB 192 KVNQEISWVRFPMPGVDVRIYHGTFTSGVILKELGFKRPGDQNKDPAERNVSK----- 243
QY 251 MNTETLSQVNPFRMFTMTNKASSNEPSLKELEK---DPVWKKLNAVKNQVLDLRLDLW 306
DB 244 ---ERISAMGDGVLFYFTFD-KGNEKKGSELEKEYINDPLFKNLNAVKNQKAYKVDDVIV 299
QY 307 ARSRGLISS-----EEMAKELVE 324
DB 300 NTAGGVMAANLLDDIEKRFVK 321

RESULT 33
Q73DM9 PRELIMINARY; PRT; 321 AA.
ID Q73DM9
AC Q73DM9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN OrderedLocusNames=BCE0683;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=22523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic

```

```

QY 132 KGIYKDLNKIAPTIELKSGFDGYNENIDAFKTIKALGKEEGKRLBEHDKIEEYKKE 191
DB 132 EKVYEQKAIAPTVPSETLGEWKDN---FKFYAKALNKEKGQKVLAAAYDKRMKDLKAK 188
QY 192 ITMDKNQKVLPAVAAKSGLLAHPNSYVQFLSQLGFK-----EALSDVTGKLSKYLKGP 247
DB 189 LGDKVQNEISWVRFPMPGVDVRIYHGTFTSGVILKELGFKRPGDQNKDPAERNVSK----- 243
QY 248 YLQMTETLSQVNPFRMFTMTNKASSNEPSLKELEK---DPVWKKLNAVKNQVLDLRL 303
DB 244 -----ERISAMGDGVLFYFTFD-KGNEKKGSELEKEYINDPLFKNLNAVKNQKAYKVDD 296
QY 304 DLWARSGLISS-----EEMAKELVE 324
DB 297 VIVNTAGGVMAANLLDDIEKRFVK 321

RESULT 31
Q6HNJ9 PRELIMINARY; PRT; 321 AA.
ID Q6HNJ9
AC Q6HNJ9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron(III) dicitrate ABC transporter, periplasmic protein.
GN Name=fecB; OrderedLocusNames=Bt9727_0525;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27."
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017355; AAF2409.1; -
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 321 AA; 36373 MW; 478614F471C1CF3 CRC64;

Query Match 23.9%; Score 404; DB 2; Length 321;
Best Local Similarity 31.4%; Pred. No. 1.9e-14;
Matches 101; Conservative 73; Mismatches 120; Indels 28; Gaps 7;

QY 16 LLFVLIIATACGNSSNSK-SKDGVEIKHEGTTKVPKPKRVVLEYSFVDALVA 74
DB 15 LLAFSLLSACGKNTKESKEDTKKEMIPVHAMGKTEVPANPKRVVILNEGTEALLE 74
QY 75 LDVPGVGIADNKNRIIPKLDKIGKVTSGTRKQPNLEISKLKPLIIADNNRHGI 134
DB 75 LGVPGVAGKSWTGDPTWPHIKDKMKVKVGDGQVNVETIASLKPDLIIQNKVHEKV 134
QY 135 YKDLNKIAPTIELKSGFDGYNENIDAFKTIKALGKEEGKRLBEHDKIEEYKKEITM 194
DB 135 YEQLKAIAPTVPSETLGEWKDN---FKFYAKALNKEKGQKVLAAAYDKRMKDLKAKLGD 191
QY 195 DKNQKVLPAVAAKSGLLAHPNSYVQFLSQLGFK-----EALSDVTGKLSKYLKGPYLO 250
DB 192 KVNQEISWVRFPMPGVDVRIYHGTFTSGVILKELGFKRPGDQNKDPAERNVSK----- 243
QY 251 MNTETLSQVNPFRMFTMTNKASSNEPSLKELEK---DPVWKKLNAVKNQVLDLRLDLW 306
DB 244 ---ERISAMGDGVLFYFTFD-KGNEKKGSELEKEYINDPLFKNLNAVKNQKAYKVDDVIV 299
QY 307 ARSRGLISS-----EEMAKELVE 324
DB 300 NTAGGVMAANLLDDIEKRFVK 321

RESULT 33
Q73DM9 PRELIMINARY; PRT; 321 AA.
ID Q73DM9
AC Q73DM9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN OrderedLocusNames=BCE0683;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=22523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic

```



```
Db 115 PDLIITASPRGAIRKNELEQIAPTVMFDPSTSNNDHFAEMTETFKQIAKAVGKEBEGKV 174
Qy 178 LEEHKKKEEYKEI-----TMDKNQKVLPAVAKS-----GLLAHPSNSVVGFLSQIGPK 229
Db 175 LADMDFADAKAKIEKADLKDKNITAMAAQAFKAKNVPTFRILT--DNSLALQVTKLGG-- 230
Qy 230 EALSDDVTGKGLSKYLKGPVLOMNTTSLQVNER-MFMTNKAASSNEPSLKELEKDPVWK 288
Db 231 --LTTTFEAKSE--PDGFKQTTVESLQSVQDSNFIYIIVADEDNIFDTQLK---GNPAWE 283
Qy 289 KLNNAVKNQKVDILDRDLWARSRLISSEEMAKELVEL 325
Db 284 ELKPKKENKMYKLKGDWTWIFG-GPESATSLATQVADV 319

RESULT 36
Q81DY8 PRELIMINARY; PRT; 320 AA.
AC Q81DY8
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE Iron(iii) dicitrate-binding protein.
GN OrderedLocusNames=BC2208;
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22608415; PubMed=1271630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Gotsman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Fonseth M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.C.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis."
RL Nature 423:87-91(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Stjerne;
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017038; AAP8458.1; -.
DR EMBL; AE017334; AAT33889.1; -.
DR EMBL; AE017225; AAT56722.1; -.
DR TIGR; BA4766; -.
DR TIGR; GBAA4766; -.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0008827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR008091; Ferrichrome_bind.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR PRINTS; PR01715; FERRIENDGPP.
KW Complete proteome.
SQ SEQUENCE 320 AA; 36173 MW; 95ED63B62B07323F CRC64;

Query Match 23.1%; Score 391; DB 2; Length 320;
Best Local Similarity 30.2%; Pred. No. 9.7e-14;
Matches 98; Conservative 64; Mismatches 138; Indels 24; Gaps 5;

Qy 9 KILSVIGLLFVLIIATAACG-NSSSSNSKSSKDGVEIKHEGTTKVPKPKRVVLEYS 67
Db 5 KKLIMFCIMLVIIAGVAGSKEEKENNISAKADSVTVKHAMGETTVNGTPPKRVVLTNE 64
Qy 68 FVDALVALDVKPGVIADNDKNRIIKPLRDKIGKTSVGRKPNLEISLKLKPDIIAD 127
Db 65 GAELLAVGTPVGTTPKPRAGDEWYHLAKELKNTVEVGTERTDINLEAVKGLKPDIIIGN 124
Qy 128 NNRHGIYKDLNKKIAPTIELKSPDGYNNENIDAFKTSIKALGKEBEGKRLKEHDKIEE 187
Db 125 KMRHEKIYDQLEIAPTVAETLRGDWKEN---FTLYTRAVNKEGQKALDDYKRIEG 181
Qy 188 YKEITMDKNQKVLPAVAKSGLLAHPNSVYQFLSQIGPK---EALSDDVTGKLSKY 243
Db 182 MKELGDKLNSKRSYIIRFVPGDVRIYQKNSFSGVNLNDIGFKRPPIQDKDDFAIKGITK- 240
Qy 244 LKGPYLQMNTEITLSQVNERMFMT-----NKASSNEPSLKELEKDPVWKLNNAVKNQV 298
Db 241 -----EQIPNWDGDLVFTTSDKQADKNNEGNSLAKEWETDPLFKQLQASKNKV 290
Qy 299 DILDRDLWARSRLISSEEMAKEL 322
Db 291 FQVDEVITWNTAGGIIKAANMLDDI 314
```

```
RESULT 37
Q81L65 PRELIMINARY; PRT; 324 AA.
AC Q81L65; Q6HSL7; Q6KLM1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN OrderedLocusNames=BA4766, BAS4424, GBAA4766;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=11392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madupu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria."
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Stjerne;
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017038; AAP8458.1; -.
DR EMBL; AE017334; AAT33889.1; -.
DR EMBL; AE017225; AAT56722.1; -.
DR TIGR; BA4766; -.
DR TIGR; GBAA4766; -.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0008827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR008091; Ferrichrome_bind.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR PRINTS; PR01715; FERRIENDGPP.
KW Complete proteome.
SQ SEQUENCE 324 AA; 36074 MW; 0EAD6197AE43B3A6 CRC64;

Query Match 23.1%; Score 391; DB 2; Length 324;
Best Local Similarity 34.4%; Pred. No. 9.8e-14;
Matches 116; Conservative 59; Mismatches 122; Indels 40; Gaps 13;

Qy 9 KILSVIGLLFVLIIATAACGNSSSNSKSSKDXD----GVEIKHEGTTKVPKPKRVVVL 64
Db 3 KILSI-----FIVVFLFVAVGCGQCKEETKADNKNQAITIKHABGETKLDKPKKVVVL 59
Qy 65 EYSFVDALVALDVKPGVIADNDKNRII---KPLRDKIGKTSVGRKPNLEISLKLK 120
Db 59 EWYSEDLLALGVQVPVGMADIKYNNKWNVTKTPSKDVV----DVGTROPNLEISRLK 114
Qy 121 PDLIITADNNRHKGIYKDLNKKIAPTIEL---KSGFDGYNNENIDAFKTSIKALGKEBEGK 177
```

```

Db      115 PDLIIITASPRGKAIKNELEQAPTVMFDPDSTNSNDHFAEMTETFKQIAKAVGKEEGKKV 174
Qy      178 LBEHDKKIEYKKEI-----TMDKQKVLPAVAAKS-----GLLAHPNSYVGQFLSOLGPK 229
Db      175 LADMDFKAFADAKAKIEKADLKDKNIAQAQFTAKNVPTFRILT--DNSLALQVTKKLG-- 230
Qy      230 EALSDDVTKGLSKYLKGPVLYQWNTETLSQVNPFR-MFIMTNKASSNEPSLELEKDPVWK 288
Db      231 --LTNTFFAKGSE--PDGFKQTTVESLQSDNSFIYIVADEDNIFDTQLK---GNPAWE 283
Qy      289 KLNNAVKNQRVDLDRDLWARSGLISSEMAKELVEL 325
Db      284 ELKFKKENQYKLGDTWIFG-GPESATSLATQADV 319

RESULT 38
Q6HCZ7 PRELIMINARY; PRT; 324 AA.
AC Q6HCZ7,
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN Names:flud; OrderedLocusts[Names-BT9727_4264];
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_Taxid=180856;
RN [1]
RN SEQUENCE 324 AA; 36067 MW; 8B2335F21181CF5F CRC64;
RC STRAIN=97-27;
RA Bretton T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL "Complete genome sequence of Bacillus thuringiensis 97-27.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS017355; AAT60882.1; -.
DR DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR DR InterPro; IPR008091; Ferrichrome_bind.
DR DR InterPro; IPR002491; Peripla_BP.
DR DR Pfam; PF01497; Peripla_BP_2; 1.
DR DR PRINTS; PR01715; FERRIBNDGPP.
KW Complete proteome.
SQ SEQUENCE 324 AA; 36067 MW; 8B2335F21181CF5F CRC64;

Query Match 23.0%; Score 389; DB 2; Length 324;
Best Local Similarity 34.1%; Pred. No. 1.3e-13;
Matches 115; Conservative 60; Mismatches 122; Indels 40; Gaps 13

Qy 9 KILSVIGLLFVLIIATAAGCNSSSSNSSKESKD---GVEIKHEEGTKVPKHPKRVVYL 64
Db 3 RIISI----FIVVLFAVCGCGQKEEKETKADNKNQAITIKHAEGETKLDKPAKKVVVL 58
Qy 65 EYGFVDALVALDQVPGVADNKKRII-----KPLRDKIKGYTSVGRKQPNLEEISKLK 120
Db 59 EWYSEDDLALGVQPVGMADIKNYKNWNTKPKSKDV----DVGTRQQPNLEEISRLK 114
Qy 121 PDLIIADNNRHKGIYKDLNKAAPTTEL---KSPGDYENIDAPKTIKALGKEEGKKR 177
Db 115 PDLIIITASPRGKAIKNELEQAPTVMFDPDSTNSNDHFAEMTETFKQIAKAVGKEEGKKV 174
Qy 178 LBEHDKKIEYKKEI-----TMDKQKVLPAVAAKS-----GLLAHPNSYVGQFLSOLGPK 229
Db 175 LADMDFKAFADAKAKIEKADLKDKNIAQAQFTAKNVPTFRILT--DNSLALQVTKKLG-- 230
Qy 230 EALSDDVTKGLSKYLKGPVLYQWNTETLSQVNPFR-MFIMTNKASSNEPSLELEKDPVWK 288
Db 231 --LTNTFFAKGSE--PDGFKQTTVESLQSDNSFIYIVADEDNIFDTQLK---GNPAWE 283
Qy 289 KLNNAVKNQRVDLDRDLWARSGLISSEMAKELVEL 325
Db 284 ELKFKKENQYKLGDTWIFG-GPESATSLATQADV 319

Db      284 ELKFKKENQYKLGDTWIFG-GPESATSLATQADV 319

```

Search completed: August 25, 2005, 09:23:53
Job time : 71 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:19:06 ; Search time 20 Seconds
(without alignments)
1611.631 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRGLKILSVIGLLFVL.....EEMAKELVELSKDKSKDK 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1232	72.9	327	D90013	hypothetical prote
2	724.5	42.9	319	E89872	hypothetical prote
3	679.5	40.2	315	B69812	ferrichrome ABC tr
4	611	36.2	348	F69831	iron(III) dicitrat
5	460	27.2	302	QRECD3	citrate-dependent
6	405	24.0	330	G89771	lipoprotein (impor
7	373	22.1	325	C69805	iron(III) dicitrat
8	350	20.7	341	E83779	iron(III) dicitrat
9	332.5	19.7	317	E69763	probable ferrichro
10	332.5	19.7	334	G84123	iron(III) dicitrat
11	327	19.3	328	AH2128	iron(III) dicitrat
12	310	18.3	325	AG2082	iron(III) dicitrat
13	295.5	17.5	330	AG2129	iron(III) dicitrat
14	292.5	17.3	330	S74458	iron(III) dicitrat
15	288.5	17.1	319	AC2421	iron(III) dicitrat
16	288	17.0	326	AF2082	iron(III) dicitrat
17	285	16.9	300	AC1855	periplasmic iron-c
18	285	16.9	342	A89846	hypothetical prote
19	284	16.8	331	AD2134	iron(III) dicitrat
20	279.5	16.5	330	AD2082	iron(III) dicitrat
21	278	16.4	361	AE2074	ferrichrome bindin
22	277.5	16.4	319	AD2085	iron(III) dicitrat
23	277.5	16.4	357	AD2079	iron(III) dicitrat
24	276.5	16.4	317	I39842	iron-uptake system
25	275	16.3	315	S32930	ferrichrome ABC tr
26	274	16.2	343	AG2075	iron(III) dicitrat
27	274	16.2	350	T36412	probable iron-side
28	273	16.2	333	AB2130	iron(III) dicitrat
29	267.5	15.8	331	AI1943	iron(III) dicitrat

30	265	15.7	313	2	B86666	hypothetical prote
31	264.5	15.7	313	2	AG1319	ferrichrome bindin
32	261.5	15.5	319	2	D70074	ABC transporter (b
33	261	15.4	330	2	H81279	enterochelin uptak
34	260	15.4	315	1	S74441	iron(III) dicitrat
35	254	15.0	341	2	B95119	hypothetical prote
36	254	15.0	341	2	F97988	hypothetical prote
37	254	15.0	349	2	AF2130	iron(III) dicitrat
38	253	15.0	321	2	G95218	hypothetical prote
39	253	15.0	321	2	E98082	hypothetical prote
40	250.5	14.8	329	2	D83778	B subtilis ferric
41	249.5	14.8	323	2	AG1378	ABC transporter, s
42	246.5	14.6	341	2	AG2880	conserved hypotet
43	246.5	14.6	341	2	G97656	iron(III) dicitrat
44	242.5	14.3	322	2	AB2078	ferrichrome bindin
45	240.5	14.2	313	2	AG1691	

ALIGNMENTS

RESULT 1

D90013
hypothetical protein SA1979 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D90013
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D90013
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <KUR>
A:Cross-references: UNIPROT:Q99S89; GB:BA000018; PID:gi3701977; PIDN:BAB43269.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1979
C:Superfamily: ferrichrome-iron transport protein fecB

Query Match	72.9%;	Score	1232;	DB 2;	Length	327;			
Best Local Similarity	71.9%;	Pred. No.	2.6e-62;						
Matches	235;	Conservative	42;	Mismatches	50;	Indels	0;	Gaps	0;
QY	5	VRGLKILSVIGLLFVL	IATAACGNNSSNSSKESKDGVEIKHEEGTTKVPKPKRVVVL	64					
Db	1	MRGLKTFISLGLI	VALLVAAACGNTDSSKESSTKDTISVKDENGTVKVPKDAKRIVVL	60					
QY	65	EYSFVDALVALDV	KPVGITADDNKNRIIKPLRDKIGKVTSGTRKQPNLBEISKLKPDLI	124					
Db	61	EYSFADALAA	LVDKPVGITADDGKKRIIKPVREKIGDVTSGTRKQPNLBEISKLKPDLI	120					
QY	125	IADNRHKGIVK	LDLNKIAPTIELKSGFDGYNENIDAFKTSKALGKEEGCKRLBEHDKK	184					
Db	121	IADSRHKGINK	ELNKIAPTILSLKSGFDGYKQNNISFNTIAKALNKEGEKRLAEHDKL	180					
QY	185	IEEYKKEITMD	KQKQVLPAAVAAKSGLLAHPNSYVQGLSQFGFKEALSDDVTKGLSKYL	244					
Db	181	INKYKDEIKF	DRNQKVLPAVAKAGLLAHPNYSYVQGLNELGFKNALSDDVTKGLSKYL	240					
QY	245	KGPVLMQNTET	LSQVNPDMFTIMTNKASSNEPSLSEKLEKDPVWKKNLNAVKNQRVDDILDRD	304					
Db	241	KGPVQLDTEH	LADLNPERMIINTDHA	KKDSAEFKQLQEDATWKKLNNAVKNRRVDIVDRD	300				
QY	305	LWASRGLISSE	MAKELVELSKSK	331					
Db	301	VWASRGLISSE	MAKELVELSKKEQ	327					

RESULT 2

Db 289 VKNNQVSDLDLDEITWNLGGYKSSKLIDDLYE--KLNIEKQSK 330

RESULT 7

C69805

iron(III) dicitrate transport permease homolog yfiy - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: C69805

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, M. A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: C69805

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-325 <KUN>

A;Cross-references: UNIPROT:O31567; GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12673.

A;Experimental source: strain 168

C;Genetics:

A;Gene: yfiy

C;Superfamily: iron(III) dicitrate transport protein

Query Match 22.1%; Score 373; DB 1; Length 325;

Best Local Similarity 29.8%; Pred. No. 4.5e-14;

Matches 96; Conservative 75; Mismatches 129; Indels 22; Gaps 8;

Qy 14 IGLFV----LTATACGNNSSSSKSSKDGVBIEHEGTT-KVPKPKRVVVLVLEYSF 68

Db 5 ISMLFVFLMAVNLASCNSSSSSSSSKTRVTKHMGTSNDIPANPKRVIVLTNEG 64

Qy 69 VDVALDVKPGVGIADNNKRIIKPLRDKIKGYTSVGRQPNLEISKUKLPDLIIADN 128

Db 65 TEALLALGKPKGVAKSKGDPWYDLKDDMKGVNGLTEPNVEAIAELKPDIIIGNK 124

Qy 129 NRHKGIYKDLANKIAPTIELKSGFDGYNENIDAFKTSKALGKEERGGKRLREHDKKIEY 188

Db 125 VRQEKIYDQLANAIAPTVFAESLAGNWKONLTLY---ANAVNKADKGKVIADFDKRVSDL 181

Qy 189 KKEITMDKKNQKVLPAVAAKSG-LLAHPNSYVVGQFLSQLGPKFKAALSDVTKGLSKYLGK- 246

Db 192 KNLG-DQNTKTVSVVRLSGESRIYVTSFPGIILDQLGFKP-----EKQVELFKKQK 235

Qy 247 ---PYLQWNTETLSQVNPFRPMTNKASS---NEPSKLEKPKDPVKKLVNAKQNRVDI 300

Db 236 DQFTFTSDSKESIPDMADVLFYFYKADNAKENENKQWNTSSSLWKNLXAVKSGNAHE 295

Qy 301 LDRDLWARRSGLISSEEMAKEL 322

Db 296 VDDVVVMTTAGGIKAANYLLDDI 317

RESULT 8

E83779

iron (III) dicitrate transport system (permease) BH1037 [imported] - Bacillus halodurans

C/Species: Bacillus halodurans

C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C/Accession: E83779

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira, Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: E83779

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-341 <STO>

A;Cross-references: UNIPROT:Q9KE21; GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA8047

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH1037

Query Match 20.7%; Score 350; DB 2; Length 341;

Best Local Similarity 29.7%; Pred. No. 9.4e-13;

Matches 102; Conservative 63; Mismatches 138; Indels 40; Gaps 11;

Qy 11 LSVGLGLFVLIIATPAACGNNSSNSS-----KSSKDGVE-----IKHEEGITKV 54

Db 6 LFTLAMSFFVLLAACGNEPSEGEPEAGDQEQVEETGDEGASGSPYSVEDVFGTVEF 65

Qy 55 PKHPKRVVLEYSFVDALVALDVKPGVGIADNNKRIIKPLRDKIKGYTSVGRQPNLE 114

Db 66 EDVPERIVLVLEWTVAEDLLALGVQPGMADLEGFEAWVDIEAELSDVVDVGTQEPNLE 125

Qy 115 EISKLPDLIIADNNRHKGIIYKDLINKIAPTIELKSPDGD-----YNENIDAFKTSKALG 169

Db 126 EIAALEPDLIIATKPRHEAIKDELIIAPTLPFPPEYDESIINOYDEWVTTFTTIAKVQ 185

Qy 170 KEEGKKRLREHDKKIEYKK---EITMDKKNQKVLPAVAAKSG-----LLAHPNSYVVG 221

Db 186 REDEAATVLAEMEQKFEAKAEIAEADLESNEFL--AMAYSGPQAPAIRFAPNGMATI 243

Qy 222 FLSGLGPKFKAALSDVTKGLSKYLPYLOMNTETLSQVNPFRPMTNKASSNEPSLKE 281

Db 244 ILENGLGVNHPDAFE-----MTG-YSEVNVLEALPEYE-DANFIYVVOEDDNVFE-DQL 295

Qy 282 EKDPVKKLVNAKQNRVDIILDRDLWARRSGLISSEEMAKELVE 324

Db 296 ADNPVWNLNVPQERTYSLGDAWLYG-GPLSAKTLVDRLVD 337

RESULT 9

E89763

probable ferrichrome ABC transporter yclQ - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: E89763

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, M. A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: E89763

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-317 <KUN>

A;Cross-references: UNIPROT:P94421; GB:Z99106; GB:AL009126; NID:G2632653; PIDN:CAB12191.

A;Experimental source: strain 168

C;Genetics:

A;Gene: yclQ

C;Superfamily: iron(III) dicitrate transport protein

Query Match 19.7%; Score 332.5; DB 1; Length 317;

Best Local Similarity 33.0%; Pred. No. 8.2e-12;

Matches 108; Conservative 59; Mismatches 119; Indels 41; Gaps 15;

[illegible]

RESULT 10
G84123
iron (III) dicitrate transport system (iron (III)-binding protein) BH3791 [imported] - E
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: G84123
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira-
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A8350; MUID:20512582; PMID:11058132
A:Accession: G84123
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <STO>
A:Cross-references: UNIPROT:Q9K6D8; GB:BA0001520; GB:BA000004; NID:g10176401; PIND:BA0707
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3791

Query Match	19.7%;	Score	332.5;	DB 2;	Length	334;			
Best Local Similarity	26.4%;	Pred. No.	8.8e-12;						
Matches	92;	Conservative	75;	Mismatches	136;	Indels	45;	Gaps	10
Qy	3	ESVRGLKTLISVIGLLFVLIIATAACGNNSSSSKSKE-----SSXGKV-EIKHEGGTTKV	54						
Db	9	QRVHSKAVLLLSALLALL--AACGGEATEGEGEEQPEATETTESGVREVSANGVVTI	66						
Qy	55	PKHPRVVVLSYSFVDALVALDVKPVGIADDNKNRIIKPLRDKIGKTVTSVTRQPNNLE	114						
Db	67	EKGPERIVTLQOGATDAVAGMTVPGVVESWLEPPMYEYLRDDEEGVEFVGQETQPNLE	126						
Qy	115	EISKLPLDIIADNNRHGIIYKDLNKIAPTIELKSFDDGNYENIDAFK-----TISKALCK	170						
Db	127	EIAKLQPLIIIASQRHEEVYDQLSEIAPT-----VHETVPEFKETVELMGAEAMDE	178						
Qy	171	EEGKGRLEEHDKKILEEVKKEITMDKNQKVLPAVAASKGLLAHPNSVYVQQR-----LSOL	226						
Db	179	EKAADIILAANDERVEDEPKTV-VEKMGDEWPEVSVLNFREDRHARIYATGFSAILLEL	237						
Qy	227	GFK-----EALSDDVTKGLSKYLKGYLYQWNTETLSQVNPENRPMFIMTNKASSNEPSLKELE	282						
Db	238	CFEPENQKNSDEI-----FMLTDMESIPEMNADVFYMFMSDDDETQKTYVEEVT	286						
Qy	283	KDPVWKKLNAVKNQRVDILDRDLWARSRLGISSEEMAKEL-----VELSK	327						
Db	287	SHPLWENLDAVKADQVYVWDEIATMNGGGGINAANLMLDDIDYDRFELEK	334						

RESULT 11

AH2128
 iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 C:Species: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AH2128
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AH2128
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-328 <KUR>
 A:Cross-references: UNIPROT:O8YTV0; GB:BA000019; PIDN:BA074282.1; PID:g17131675; GSPDB:C
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr2583
 C:Superfamily: ferrichrome-iron transport protein fecB

Query Match	19.3%	Score	327;	DB	2;	Length	328;		
Best Local Similarity	29.0%	Pred. No.	1.7e-11;						
Matches	97;	Conservative	71;	Mismatches	126;	Indels	32;	Gaps	12;
Qy	8	LKILSVIGLLFVLIIATAACGNSSNSSKSSKDGVEIKHEEGTTKVPKHPKRVVLEYS	67						
Db	26	LILFSIITALIVICEMSTNNVTIINSVATSEMRV-VKHTMETKIPLRPQRVVL--G	82						
Qy	68	FVDALVALDVKPGV---IADNKNRIIRKIDRDKIGYTSVGRKQPNLSEIKLKPDLI	124						
Db	83	GLDNILALGVKPIAATTFSSDN---PADYLQDVTSIEKIGINGQPNLEKILYKPDLI	138						
Qy	125	IADNRHKGYYKDLNKIAPTIELKSPGDGVNENIDAPKTIISKALGKEEGKRLREHDKK	184						
Db	139	LG-FSMDAELYGQLSQIAPT-V-LADQSDWK---DWLKKYAEAVGETAKAKLLAQEYDOR	193						
Qy	185	IEEYKKEI--TMDKNQKVLPAVAAKSGLLAHPNSYVQGLFSQLGKFEALSDDVTKGLSK	242						
Db	194	MESLRQMGDTLSQTKVSLVNFWANFTRL--YMNNSFGGSLKEIGLPRPKYQDKDKN---	249						
Qy	243	YLKGPYLMQNTETLSQVNPERMITMTNKASNPESLKELEKDPVWKIKLNAVKNORVDIL	302						
Db	250	-----HENISLELIPQINGDVI FLIL--CGHNESRLKQFTNHPLWSQLQAVQKNQVYPVT	302						
Qy	303	RDWLWARSRLGISS-----EEMAKELVE	324						
Db	303	GDWVISAWGIIGANILVLDLFFKYILE	328						

RESULT 12

AG2082
iron(III) dicitrate-binding periplasmic protein of ABC transporter alr2213 [imported] -
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AG2082
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, T.
DNA Res. 8, 203-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An-
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2082
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <NR>
A:Cross-references: UNIPROT:Q8YUW8; GB:BA000019; PIDN:BAB73912.1; PID:gl7131304; GSPD8:
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2213

C;Superfamily: ferrichrome-iron transport protein fecB

Query Match 18.3%; Score 310; DB 2; Length 325;
Best Local Similarity 28.3%; Pred. No. 1.6e-10;
Matches 89; Conservative 71; Mismatches 124; Indels 30; Gaps 8;

Qy 6 RGLKILSVIGLLFVLIIATAACGNSSNS-----SKESKDGVEIKHEGGTTKVPKPKRV 61
Db 7 RYIKFLILPLTLVIV--NSCIYINPENSGLITRIKTSBCRLIKHPGSCVPVKQVR 64
Qy 62 VLEYSFVDALVADVPVGIADONKKNRIIKPLDKIGKIVTSVGTQRQPNLEISKLKP 121
Db 65 IALDETSWEALLADLPKPIATAQPNIAAGSIQKLGKABGIVSLGKQOPNIEKMWQLNP 124
Qy 122 DLIADNNRHGIYKDLNWKIAPTIELKSFQDGYNENI--DAFKTISKALGKEEGKKRLE 179
Db 125 DLILGFSVSAEQ-VKLFSQLAPTIVTL-----DYIQFAWKDALSRIAIIDKSEQAKILLE 178
Qy 180 EHDKKIEYKEITMDKNQKVLPAVAKSGLLAHPNSNSYVGQFSLQSLGFEKALSDVV--- 236
Db 179 QYQQRVKELRTFINYNLKEKTVSVSRFYAG-----NQVPEFRTKYSFPGLLTVRIP 231
Qy 237 ----TKGLSKYLKGYLQMTETILSQVNPFRMFMTNKAASNEPSLKELEKDPVWKILNA 292
Db 232 VPENQNLTTNENQPLVSVLSLERLELDADVLFVALDPGA--ESFQKYQNTPLWQKLDV 289
Qy 293 VKNQRVDILRDLM 306
Db 290 VKNQRVYAVDSGYW 303

RESULT 13

AG2129
iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AG2129
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-330 <KUR>
A;Cross-references: UNIPROT:O8YTX3; GB:BA000019; PIDN:BA74289.1; PID:gl7131683; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr2590

Query Match 17.5%; Score 295.5; DB 2; Length 330;
Best Local Similarity 27.6%; Pred. No. 1e-09;
Matches 93; Conservative 77; Mismatches 130; Indels 37; Gaps 10;

Qy 8 LKILSVIGLLFVLIIATAACGNSSNSKESKDGVE----IKHEGGTTKVPKPKRVV 63
Db 10 LLLLSVCTFLV-----TGCNSNVSKLPQNTTSNITTCRCVINHEAGKTQICGQPKKVA 65
Qy 64 LEYSFVDALVADVPVGIAD-----DNKNRIIKPLADKI-GKIVTSVGTQRQPNL 113
Db 66 LSPVPLDMMALGVGPAGVAYEVDLLNSKVFORPKEQ-IPYLDGDRITSQPMNVGDRGNPSL 124
Qy 114 BEISKLPDIIADNNRHGIYKDLNWKIAPTIELKSFQDGYNENIDAFKTSKALGKEE 173
Db 125 ESLQLKPDILGASYNETEQIILNKIAPTILISHRDGITPAWQQTISIQAALGREDK 184
Qy 174 GKRLREHDKKIEYKKEIT-MDKNQKVLPAVAKSGL-----LAHPNSNSYVGQFSLQSLG 228
Db 185 VPTVIAEYQKLSQAKTALPTAQHQELL--LLAFRGITLTSTFTGAETFFAGLLQDLGF 242
Qy 229 KEALSDDDVTGKLSKYLKGYLQMTETILSQVNPFRMFMTNKAASNEPSLKELEKDPVWK 288

Db 243 K-LVSPESTE-----YEVALEVLPKFSGDLIVMPSGNNNIENAKRQWQNPLIQ 291
Qy 289 KLNKVNQRVDILDRDLWARSRGLISSEEMAKELVEL 325
Db 292 SISAHTNRIYFIDYQLASRIRGPIAAELFVNQVRQL 328
RESULT 14
S74458
iron(III) dicitrate-binding periplasmic protein fecB-2a - Synechocystis sp. (strain PCC N;Alternate names: protein slr1491
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Accession: S74458
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Yasuda, M.; Yasuda, M. DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74458
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-330 <KAN>
A;Cross-references: UNIPROT:P72610; EMBL:D90899; GB:AB001339; NID:gl651650; PIDN:BAAL661
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: fecB 2a
C;Superfamily: iron(III) dicitrate transport protein

Query Match 17.3%; Score 292.5; DB 2; Length 330;
Best Local Similarity 28.4%; Pred. No. 1.5e-09;
Matches 85; Conservative 68; Mismatches 119; Indels 27; Gaps 9;
Qy 11 LSVIGLLFVLIIATAACGNSSN---SSKESKDGVEIKHEGGTTKVPKPKRVVVLVYS 67
Db 13 LTILITIVFF-----SACVGSSTQNLDQSTELLSVDCRIVEHSLGKTCVPLEPRRVVALDGA 68
Qy 68 FVDALVADVPVGIADONKKNRIIKPLDKIGKIVTSVGTQRQPNLEISKLKPDIIAD 127
Db 69 TVGNLLALGMPAGVA-----SNLLPEITRLIPNVPRLGQSSQINLETALVLQPDIIIGA 123
Qy 128 NNRHKGIVKDLNWKIAPTIELK-SFGDGYNENIDAFKTSKALGKEEGKKLEBHDKKIE 186
Db 124 VWEMKGIYNKLSAIAPTVAFEMQTPADWQR---PFRFDGQVLGLETAQAEKLEQYQMRVN 180
Qy 187 EYKKEITMDKNQKVLPAVAKSGLLA-HPSNSYVGQFSLQSLGFEKALSDDV-TKGLSKYL 244
Db 181 KLRQVSDSPQLQISLVRIRAESGQNSLYLKNCFGAILADLGFARPPSDQGTDPQPPFA 240
Qy 245 KGPYLQMTETILSQVNPFRMFIMT-----NKASSNEPSLKELEKDPVWKILNAVKNQRV 298
Db 241 K-----SISRESMEADGDVIFLFTFGHTPPQIAAAAEQAQLERLDTDPLWQSLGAVQKNRV 295

RESULT 15

AC2421
iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AC2421
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <KUR>

С.: ШИЛДА, И.; НАСЛОВИ, И.; СУВАВАГА, И.; НАЈУДИ, И.; НАСТАВА, И.

Db 245 GGYTFGLETHPMDADGDVITSHWDEDEGAKALKRNQNNPLWRKLKAVEQNQVYVNRH 305

QY 305 LWASRGRGLISSEEMAKEL 322

Db 305 FW-RGSNTLAAHKVTDL 321

RESULT 21

Az2074

ferrichrome binding protein of ABC transporter all2147 [imported] - Nostoc sp.

Db 294 YFEPQMKLQKQMSNAIAQSLTASKEGRVYFIPAYLCGLGPGPIGTLYLQELQKQL 352

RESULT 24
I39842
iron-uptake system (binding protein) feua - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C/Accession: I39842; E69621
Biochim. Biophys. Acta 1186, 27-34, 1994
A/Title: Isolation of Tn917 insertion mutants of Bacillus subtilis that are resistant
A/Reference number: I39839; MUID:94281248; PMID:8011666
A/Accession: I39842
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-317 <RES>
A/Cross-references: UNIPROT:P40409; GB:L19954; NID:9438454; PIDN:AAA64354.1; PID:9438458
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Poulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, E.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: E69621
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Residues: 1-317 <UN>
A/Molecule type: DNA
A/Cross-references: GB:Z99104; GB:AL009126; NID:92632267; PIDN:CAB11939.1; PID:g2632430
A/Experimental source: strain 168
C/Genetics:
A/Gene: feua
C/Superfamily: iron(III) dicitrate transport protein

Query Match 16.4%; Score 276.5; DB 2; Length 317;
Best Local Similarity 26.5%; Pred. NO. 1.2e-08;
Matches 86; Conservative 72; Mismatches 136; Indels 31; Gaps 12;
Qy 9 KILSVIGLLFVLIATAACGNSSSSK-----ESSKGVETKHEGTTKPKPKRVVVL 64
Db 3 KISUTLLILLALTAAACGSKNASTASKAGTASEKKIEYLDKTYEVTVPTDKIAITGS 62
Qy 65 EYFVDALVALDVKPVG-IADNKKRIKPLRDKIGKYSVGTGRKQPNLEISKLKPD 123
Db 63 VESMEDAKL-LDVHPQGAISFGKPPDMFKDITDKA---EPTGEKWEFNIEKLEMKPDV 118
Qy 124 IADNNRHGIYKOLNKIAPTIELSKFGDYNENIDAFKTSKALGKEEGKRLSEHDK 183
Db 119 ILASTKFPPEKTLQKISTAGTTPVSHISSNKKENN---MLLAQLTGKKEKAKKIADYE 175
Qy 184 KIEEYKKEIT-WDKNQKVLPAVAASKGLLAHPNSYVQFL-SQLGPKREALSDVTKGLS 241
Db 176 DLKEIKTKINDAKDSKALVIRIRGNIIYIPEQVFNSTLYGDLGLK-APNEVAAKA 233
Qy 242 KYLGPYQLMNTETLSQVNPFRMFI-MTNKASSNEP-SLKELEKDPVWKKLNAVKNQRVD 299
Db 234 QELS-----SLEKLENNPDHIFVQPSDDENADKPDALKOLEKNPIWLSLKAVKEDHY 287
Qy 300 ILDRD-----LWARSGLISSEE 317
Db 288 VNSVDPLAAGGTANSKVRFLKAAAE 312

RESULT 25
S32930
ferrichrome ABC transporter fhud precursor - Bacillus subtilis
N/Alternate names: ferrichrome-binding protein fhud
C/Species: Bacillus subtilis
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
A/Accession: S32930; C69622
R/Schneider, R.; Hantke, K.
Mol. Microbiol. 8, 111-121, 1993
A/Title: Iron-hydroxamate uptake systems in Bacillus subtilis: identification of a lipop
A/Reference number: S32930; MUID:93268086; PMID:8388528
A/Accession: S32930
A/Molecule type: DNA
A/Residues: 1-315 <SCH>
A/Cross-references: UNIPROT:P37580; GB:M87283; NID:9289271; PIDN:AAA22424.1; PID:g289272
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Poulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, E.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69622
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Residues: 1-315 <KUN>
A/Molecule type: DNA
A/Cross-references: GB:Z99121; GB:AL009136; NID:g2635827; PIDN:CAB15338.1; PID:g2635846
A/Experimental source: strain 168
C/Genetics:
A/Gene: fhud
C/Superfamily: Bacillus subtilis ferrichrome ABC transporter fhud
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-315/Product: fhud protein #status predicted <MAT>

Query Match 16.3%; Score 275; DB 2; Length 315;
Best Local Similarity 32.1%; Pred. No. 1.4e-08;
Matches 97; Conservative 42; Mismatches 121; Indels 42; Gaps 13;
Qy 17 LFLVIATAACGNSSSSSKESKDGVE---IKHEEGTTKVPKPKRVVVLVEYSFVDALV 73
Db 15 LLLIALAAGCNSESKEGASDSK-GAETFTYKAENGKVKIPKPKRVVWADGYGYPK 73
Qy 74 ALDVKPGIADNKKRIIKPLR---DKIGKYSVGTGRKQPNLEISKLKPDLLIADNRR 130
Db 74 TLGINVVGAPENVFKNPYYKGTNGVENIGDGTSV-----EKVIDLNPDLIIVWTQ 125
Qy 131 HKGIYKOLNKIAPTIELSKFGDYNENIDAFKTSKALGKEEGKRLSEHDKKEEYKK 190
Db 126 GADI-KKLEKIAPIVAVKY---DKLDNIEQLKEFAKMTGTDKAEKWLAKDKRVAAK- 180
Qy 191 BITMDKQKVLPAVAASKGLLAHPNSYVQFLSQLG-----FKREALSDVTKGLSKYL 244
Db 181 -----TKIKKAVGDKTISIMQTKGDIYVFKDPGRGSGSIYKD-LGLQATK-LTKRK 231
Qy 245 ---KGP-YLQNTETLSQVNPFRMFI-MTNKASSNEP-SLKELEKDPVWKKLNAVKNQRVDI 300
Db 232 AIDQPGYTSISLEKLPDFAGDYIFAGPWQSGDGGVFE---SSIWKNLNAVKNGHVYK 288
Qy 301 LD 302
Db 289 MD 290

RESULT 26

AG2075
iron(III) dicitrate-binding periplasmic protein of ABC transporter all2157 [imported] -
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AG2075
R:Kanezo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana-
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <KUR>
A:Cross-references: UNIPROT:Q8YV24; GB:BA000019; PIDN:BA073856.1; PID:gl7131248; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
C:Gene: all2157
C:Superfamily: ferrichrome-iron transport protein fecB
Query Match 16.2%; Score 274; DB 2; Length 343;
Best Local Similarity 26.5%; Pred. No. 1.8e-08;
Matches 91; Conservative 80; Mismatches 122; Indels 50; Gaps 14;
QY 6 RGLKILS-----VIGLFLVLIATAACGNSS-----SNSKSSKDGVEIKHEGTTKV 54
DB 17 KSMKIISRRFVFLFLGILIFTAIWVSSKSFYEATNSIPQONAKCRV-VQHVQGETCI 75
QY 55 PKHPKRVVLEYSFVDALVALDVKPGV-----IADD-----NKNRIIPLRDKIGKYS 104
DB 76 PLKQRIIVLDFNSFAALLALDTPKIATWITTEIDDFYFOCKAGVILRSSG----- 131
QY 105 VGTROPNLEESKLPDLIIA-DNNRHGIYKDLNKAIPTELKSF-DGDYENIDAF 161
DB 132 -----QINLEKLVLPDIIIVISHPFGAGIYKASQIAPTIVLPWIETRGWKKH---- 183
QY 162 KTISKALGKEEGKLEEDHKKIEYKKEITWMDKNQ-KVLPAVAASKGLLAHPSNSYVG 220
DB 184 QOTARLNNRRETSIQIYNNYQNRVQKQLSNNHKKIHVSFAVAGQLVITRQKSPAG 243
QY 221 QFLSQIGFKEALSDVTKGLSKYLKGPY-LQWNTETLSQVNPMPIMTNKASSNEPSLK 279
DB 244 GILHDIGILNPFAE-----SGDYLPLESELLPNIDSDILFIAPLR-KDDYSVIK 293
QY 280 ELEKDPVWKLNAVKNQVRVDILDRDLWARSGLISSENAKEL 322
DB 294 KLQKPLWKLKRAVQOQNYIVDFSVW-RGLNMLAAYAMLDL 335
RESULT 27
T36412
probable iron-siderophore binding lipoprotein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36412
R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21606
A:Accession: T36412
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-350 <SAU>
A:Cross-references: UNIPROT:Q9RK12; EMBL:AL109974; PIDN:CAB53324.1; GSPDB:GN00070; SCOE
A:Experimental source: strain A3(2)
C:Genetics:
C:Gene: SCOEDB:SCF34.13C
C:Superfamily: iron(III) dicitrate transport protein
Query Match 16.2%; Score 274; DB 2; Length 350;
Best Local Similarity 29.9%; Pred. No. 1.8e-08;
Matches 95; Conservative 56; Mismatches 113; Indels 54; Gaps 14;

QY 14 IGLFLVLIATAACGNSSSSSKSSKDG-----VEIKHEGTTKVPKPKRVVVL 64
DB 25 VGLL-----AGCG--SDSDPADPAGGCTPAAGAFPTVEHAFGTTIKDKAPRVVSV 76
QY 65 EYSFVDALVALDVKPGVGIADNKKRIIKP-----LRDKIGKYSVGTGRKQ-----PN 112
DB 77 GTTDDQTVLAFGIKPGVGVQVQPNPAGQSPDINTQWPVKDKMGDTKPEVVMKNGDTGN 136
QY 113 LEEISKLKPLDIADNNR-HKGIYKDLNKAIPTI-----ELKSPDGDYNNENIDAFKTSK 166
DB 137 FEKIAALRPDLIVAYSEIDQAAVDKLSKIAPTGTGRTGKEKPPSAPWQDNA---LHIAK 193
QY 167 ALGKEEGKRLLEEDHKKIEYKKEIT--MDKNQKVLPAVAASKGLLAHPSNSYVQFPLS 224
DB 194 ALGKAEEGKRVADIQGKLDAKQHPFADKTAIVLSWY--KDSVAPPTSTDVGRGLVT 251
QY 225 QLGPK-EALSDDVTKGLSKYLKGPYLOMNTETLSQVNPMPIMTNKASSNEPSLKELSK 283
DB 252 GIGFKYTEIDKVG-----DFYTLSPERVDLVDRVFVINDKA--DQDALKKPF- 302
QY 284 DPVWKLNAVKNQVRVDIL 301
DB 303 --LFTNLDAVKNKVSYL 318
RESULT 28
AB2130
iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AB2130
R:Kanezo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An-
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2130
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <KUR>
A:Cross-references: UNIPROT:Q8YTX0; GB:BA000019; PIDN:BA074292.1; PID:gl7131686; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
C:Gene: alr2593
C:Superfamily: ferrichrome-iron transport protein fecB
Query Match 16.2%; Score 273; DB 2; Length 333;
Best Local Similarity 28.3%; Pred. No. 1.9e-08;
Matches 90; Conservative 65; Mismatches 127; Indels 36; Gaps 15;
QY 8 LKILSVIGLFLVLIATAACGNSSSSSKSSKDGVE-----IKHEGTTKVPKPKRVV 63
DB 9 LQIFITLTVFVLV--TGCYSHPNQIQSANKPVTSECRILKPLGETCIPHPORIITA 66
QY 64 LEYSFVDALVALDVKPGVIA--DDNKNRIIPLRDKIGKYS--VGTGRQPNLEISKL 119
DB 67 MDEDILEITLVALDLKPIAVAINFDEWSSR-EKQLWQKAGSIDSVIGNHGILNLEKMLLL 125
QY 120 KPDLI--TADNNRHGIYKDLNKAIPTELKSFQDGDYNNENI--DAFKTISKALGKEEGK 175
DB 126 KPDILGLAESTDRKS-YELFSQIAPTIVV-----DYAQTAWRDVLLRVGNIGTKTEQAO 179
QY 176 KLELEHDHKKIEYKKEITWMDKNQKVLPAVAASK--GLLAHPSN-SYVQFPLSQLG--FK 229
DB 180 KLIAEFQQRKEIKRLVIVKNKLGKTKLISVVRVNYFHKTIEPRSNFSPGSLVVELGLSP 239
QY 230 EALSDDVTK-GLSKYLKGPYLOMNTETLSQVNPMPIMTNKASSNEPSLKELKDPVWK 288
DB 240 EKQNKIPTSPPG-----PFVFPASLERLDLADVMFVTLT--AGGEENFKKFPQASPLWQ 291
QY 289 KLNNAVKNQVRVDILDRDLW 306
|| ||||| | | |

[illegible]

RESULT 35
B95119
hypothetical protein SP1032 [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: B95119
R/Tettalin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unyav, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A./Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A./Reference number: A95000; MUID:21357209; PMID:11463916
A./Accession: B95119
A./Status: preliminary
A./Molecule type: DNA
A./Residues: 1-341 <KUR>
A./Cross-references: UNIPROT:Q97R09; UNIPROT:Q8DDY6; GB:AE005672; PID:AAK75147.1; PID:q1
A./Experimental source: strain TIGR4
C./Genetics:
A./Gene: SP1032

	Query Match	15.0%;	Score 254;	DB 2;	Length 341;
	Best Local Similarity	28.9%;	Pred. No. 2.3e-07;		
	Matches	93;	Conservative	58;	Mismatches 115; Indels 56; Gaps 15;
Qy	13 VIGLLFVLIATAACGNNS--- <td>69</td> <td>:</td> <td>:</td> <td>:</td>	69	:	:	:
Dd	: : : : : :	:	:	:	:
Qy	7 LIATLAWCIVFSACSSSVRNEENTSEHAPDKIVLDHPAGQTILDDKKPERVATIAMGNH	66	:	:	:
Dd	:	:	:	:	:
Qy	70 DALVALDVKPGVIADDN---KKNRIIKP-----LRDKIGKYTVSGTRKQPQNLEIRISKLP	121	:	:	:
Dd	:	:	:	:	:
Qy	67 DVALALGTVPFGSKANYGVSADGVLPWTTEKI KELNGKANLFDDLGLMFEALSINSKP	126	:	:	:
Dd	:	:	:	:	:
Qy	122 DLIIADNNRHKGI----YKDLNKIAPTTELKSFDFGDYNENI--DAFKTITSXALGKEEBE-	174	:	:	:
Dd	:	:	:	:	:
Qy	127 DVILLG---YGGTIKEDYDTLSKIAPVAAYAKSPW---QTWRDMWKIDSLAALGWMEKGD	180	:	:	:
Dd	:	:	:	:	:
Qy	175 -----KKRLLEEHDKKILEYYKETITMDKNQKVLPAV---AAKSGLAHPSNYSYVGQ	221	:	:	:
Dd	:	:	:	:	:
Qy	181 ELIKNTEARISKELEKH-----PEIKGKI---KGKKVLFTMINAADTSKFWIYTSKDPAN	233	:	:	:
Dd	:	:	:	:	:
Qy	222 FLSQLG--FKREALSDVTYKLGSKYLKGPYLQMNTETLSQVNPERFMFTMNKASNEPSLK	279	:	:	:
Dd	:	:	:	:	:
Qy	234 YLTDLGLVFPSSLKFEFSSE---DSFAK---BISAEEKINKINDADVII---TYGDDKTLE	283	:	:	:
Dd	:	:	:	:	:

Qy 280 ELEKDPVWKLLNAVNQRVDIL 301
 :|||: |:|||: ||:
Db 284 ALQKDPLLGNAINKGAVAVI 305

RESULT 36

F97988
hypothetical protein ABC-SBP [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: F97988
E:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff,
e, R.; Leblang, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAdams,
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas,
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; PMID:21429245; PMID:11544234
A:Accession: F97988
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-341 <Kur>
A;Cross-references: UNIPROT:Q97R09; UNIPROT:Q8DPV6; GB:AE007317; PIDN:AAK99738
C:Genetics:
A:Gene: ABC-SBP

Query Match	15.0%;	Score 254;	DB 2;	Length 341;
Best Local Similarity	28.9%;	Pred. No. 2.3e-07;		
Matches	93;	Conservative 58;	Mismatches 115;	Indels 56; Gaps 15;
Qy	13	VIGLLFVLIIATAACGNNSSNSSTKESKDGVEIKHEBGTIKVPHKPRVVVLEVSFV	69	
Db	7	LIATLACMIVFSACSSNVKNEENTSEKHAADKIVLDHAFQQTILDKKPERVATIAMGNH	66	
Qy	70	DALVALDVKPGIADDNKKRIKPKLRDKIKIGKYTVGTRKQPNLEISKLP	121	
Db	67	DVALAGIVPGFSKANYGSADKGLVLPWTEEKIKELNGKANLFDDLDGLNFSAISNSKP	126	
Qy	122	DLIIADNNRHKGIYKOLNKIAPTIELKSFQGDYNEIMDAFKTISKALGKEEBEG	174	
Db	127	DVILAGYSGITKEDYDLSKIAPVAAYKSKPWQTLWRDWIKIDSALGMEKGECD	180	
Qy	175	KKRLBEHDKKIEBKKEITMDKNQKILPAVAAKGGLLAHPSSNVYVGQ	221	
Db	181	ELIKNTEARISKELKHPKGIKGGKVLFTMINAADTSKFWIYTSKDPKAN	233	
Qy	222	FLSQGFKKALGDDVTGKLSKYLKGPYLOMNTETLSQVNPERFMINTKASNEPSLK	279	
Db	234	YLPDLGLVFPESLKEFSEDSPAKEISEEANKINDADVIIITYGDDKTLE	283	

Qy 280 ELEKDPVWKKLNAVKNQVDIL 301
-|:-|-:-|-:-|-:-|-:-|-:-|-:-|
Dd 284 ALOKDPLLGKINAIKNGAVAVI 305

RESULT 37
AF2130
iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2130
K:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi,
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840

A;Residues: 1-349 <KUR>
A;Cross-references: UNIPROT:Q8YTW6; GB:BA000019; PIDN:BAB74296.1; PID:g17131690; GSPDB:C

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83778
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-329 <STO>
A;Cross-references: UNIPROT:O9KE30; GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA047
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1028

Query Match		14.84;	Score 250.5;	DB 2;	Length 329;
Best Local Similarity		25.64;	Pred. No. 3.5e-07;		
Matches		86;	Conservative 71;	Mismatches 134;	Indels 45; Gaps 13;
Qy	10	ILSVIGLLFVLIATAACNNSSN-----SSKESKDGVE-----IKHEEGTT	52		
Db	11	VVGIVSLMW-----ACGADDTAEPADAEQEAAQDTVEEEMVETEAGGEVTVQHDLGEA	65		
Qy	53	KVPKHPKRVVVLVSFVDALVALDVPVGIADDNKNRIIKPLRDKIGKITYSGTRKQPN	112		
Db	66	VIPKNPTVVVDFGTLDLRLGVEVAGVPQANVPAYLSQYBEE---TYANVGTLPED	122		
Qy	113	LEEISKLPDLIIADNNRHKGIYKDLNKIAPTIELKSPDGDYNNEN-IDAPK---TISKA	167		
Db	123	PETIYDQPDIIII-SGRTSEAYDELSDIAPTL----FVGLDTENYLESFRSNMETLGEI	177		
Qy	168	LGKEEGKKRLLEHDKIEEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYVQGFLSQLG	227		
Db	178	FGKETEVVEEALASLEASTEEV-QQLAEKKEETGLIVLANDGNVSAYGPGSRFGVIHDEFG	236		
Qy	228	FKEALSDDVTKGLSKYLKGPYLQWNTETLSQVNPERMFTMTNKASSNEPSLKELEKDPVW	287		
Db	237	V-----TPVDENIEVSNHGQ--SISFEYIVKPNPHLFVIDRGAVVQGEETNIENELV-	288		
Qy	288	KKLNAVKNQRVDIILDRDLWARS-RGLISSEEMAKEL	322		
Db	289	QQTAYEGNIHYLTPTDIWYISGGGLVSVEQMIDEM	324		

Search completed: August 25, 2005, 09:27:10
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:27:00 ; Search time 163 Seconds
(without alignments)
807.548 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRGLKILSVIGLLFVL.....EEMAKVELSKDKSKDNK 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues
Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/FCRUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1690	100.0	335	US-10-724-972A-6352	Sequence 6352, Ap
2	1668	98.7	331	US-10-282-122A-71138	Sequence 71138, A
3	1242	73.5	337	US-09-815-242-12254	Sequence 12254, A
4	1232	72.9	337	US-10-282-122A-43828	Sequence 43828, A
5	1181	69.9	309	US-09-815-242-5459	Sequence 5459, Ap
6	726	43.0	319	US-09-815-242-5249	Sequence 5249, Ap
7	726	43.0	319	US-09-815-242-12122	Sequence 12122, A
8	724.5	42.9	319	US-10-282-122A-41133	Sequence 41133, A
9	724.5	42.9	319	US-10-470-048B-329	Sequence 329, App
10	644	38.1	175	US-10-282-122A-71291	Sequence 71291, A
11	516.5	30.6	298	US-10-282-122A-66890	Sequence 66890, A

12	460	27.2	300	15	US-10-282-122A-56054	Sequence 56054, A
13	460	27.2	302	9	US-09-815-242-10469	Sequence 10469, A
14	460	27.2	302	15	US-10-282-122A-56805	Sequence 56805, A
15	454	26.9	297	15	US-10-282-122A-69420	Sequence 69420, A
16	405	24.0	330	8	US-08-781-986A-5193	Sequence 5193, Ap
17	405	24.0	330	14	US-10-278-946-16	Sequence 16, Appl
18	405	24.0	330	15	US-10-329-624-5193	Sequence 5193, Ap
19	405	24.0	330	17	US-10-470-048B-346	Sequence 346, App
20	405	24.0	330	17	US-10-967-189-16	Sequence 16, Appl
21	391	23.1	324	15	US-10-282-122A-45599	Sequence 45599, A
22	331	19.6	315	9	US-09-738-626-3875	Sequence 3875, Ap
23	327	19.3	321	15	US-10-282-122A-53267	Sequence 53267, A
24	314	18.5	314	15	US-10-282-122A-46283	Sequence 46283, A
25	301	17.8	314	15	US-10-282-122A-45385	Sequence 45385, A
26	300	17.8	314	15	US-10-282-122A-46193	Sequence 46193, A
27	296.5	17.5	351	18	US-10-724-972A-7181	Sequence 7181, Ap
28	296	17.5	342	15	US-10-282-122A-68785	Sequence 68785, A
29	283	16.7	317	9	US-09-071-035-38	Sequence 38, Appl
30	283	16.7	317	14	US-10-206-576-38	Sequence 38, Appl
31	283	16.7	317	17	US-10-912-362-38	Sequence 38, Appl
32	279	16.5	317	9	US-09-815-242-10905	Sequence 10905, A
33	274.5	16.2	315	15	US-10-282-122A-46179	Sequence 46179, A
34	268	15.9	296	9	US-09-071-035-40	Sequence 40, Appl
35	268	15.9	296	14	US-10-206-576-40	Sequence 40, Appl
36	268	15.9	296	17	US-10-912-362-40	Sequence 40, Appl
37	266	15.7	306	9	US-09-738-626-4403	Sequence 4403, Ap
38	266	15.7	306	15	US-10-627-476-572	Sequence 572, App
39	257.5	15.2	342	15	US-10-282-122A-47311	Sequence 47311, A
40	254	15.0	341	16	US-10-474-776-663	Sequence 663, App
41	254	15.0	341	16	US-10-415-478A-12	Sequence 12, Appl
42	254	15.0	341	17	US-10-472-928-2046	Sequence 2046, Ap
43	253	15.0	321	9	US-09-815-242-13454	Sequence 13454, A
44	253	15.0	321	16	US-10-415-478A-9	Sequence 9, Appl
45	253	15.0	321	17	US-10-472-928-3884	Sequence 3884, Ap

ALIGNMENTS

RESULT 1
US-10-724-972A-6352
; Sequence 6352, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
; APPLICANT: Douchette-Stamm, Lynn
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT FILING DATE: 2003-12-01
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/450,969
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 6352
; LENGTH: 335
; TYPE: PRT
; ORGANISM: S.epidermidis
US-10-724-972A-6352

Query Match 100.0%; Score 1690; DB 18; Length 335;
Best Local Similarity 100.0%; Pred. No. 6.8e-107;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVESVRGLKILSVIGLLFVLITACGNNSSNSSKESKDGVEIKHEGTTKVPKPKR 60
DB 1 GVESVRGLKILSVIGLLFVLITACGNNSSNSSKESKDGVEIKHEGTTKVPKPKR 60

Qy 61 VVLEYSFVDALVDKVPVGIADNDKQRIIKPLRDKIGKYTSVGTTRKQPNLEISKLK 120
Db 61 VVLEYSFVDALVDKVPVGIADNDKQRIIKPLRDKIGKYTSVGTTRKQPNLEISKLK 120
Qy 121 PDLIADNNRHGKYKDLNKIAPTIELKSPFGDYNENIDAFKTSKALGKEEGKRLKEE 180
Db 121 PDLIADNNRHGKYKDLNKIAPTIELKSPFGDYNENIDAFKTSKALGKEEGKRLKEE 180
Qy 181 HDKIEEYKKEITWDMQKQVLPVAAKSGLLAHPNSVVGQFLSQGPKALSDDVTKGL 240
Db 181 HDKIEEYKKEITWDMQKQVLPVAAKSGLLAHPNSVVGQFLSQGPKALSDDVTKGL 240
Qy 241 SKYLKGPYQWNTETLSQVNPFRPMTNKASSNEPSLKELEKDPVWKKLNAVKNQRVDI 300
Db 241 SKYLKGPYQWNTETLSQVNPFRPMTNKASSNEPSLKELEKDPVWKKLNAVKNQRVDI 300
Qy 301 LDRLWARSGLISSEEMAKELVELSKDKSKDNK 335
Db 301 LDRLWARSGLISSEEMAKELVELSKDKSKDNK 335

RESULT 2

US-10-282-122A-71138
; Sequence 71138, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71138
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-71138

Query Match 98.7%; Score 1668; DB 15; Length 331;

Best Local Similarity 99.7%; Pred. No. 2.le-105;
Matches 330; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 5 VRGLKILSVIGLLPVLIIATAACGNSSSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVL 64
Db 1 MRGLKILSVIGLLPVLIIATAACGNSSSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVL 60
Qy 65 EYSFVDALVDKVPVGIADNDKQRIIKPLRDKIGKYTSVGTTRKQPNLEISKLKPDLI 124
Db 61 EYSFVDALVDKVPVGIADNDKQRIIKPLRDKIGKYTSVGTTRKQPNLEISKLKPDLI 120
Qy 125 IADNNRHGKYKDLNKIAPTIELKSPFGDYNENIDAFKTSKALGKEEGKRLKEEHDKK 184
Db 121 IADNNRHGKYKDLNKIAPTIELKSPFGDYNENIDAFKTSKALGKEEGKRLKEEHDKK 180
Qy 185 IEYKKEITWDMQKQVLPVAAKSGLLAHPNSVVGQFLSQGPKALSDDVTKGLSKYL 244
Db 181 IEYKKEITWDMQKQVLPVAAKSGLLAHPNSVVGQFLSQGPKALSDDVTKGLSKYL 240
Qy 245 KGPYQWNTETLSQVNPFRPMTNKASSNEPSLKELEKDPVWKKLNAVKNQRVDILDRD 304
Db 241 KGPYQWNTETLSQVNPFRPMTNKASSNEPSLKELEKDPVWKKLNAVKNQRVDILDRD 300
Qy 305 LWARSGLISSEEMAKELVELSKDKSKDNK 335
Db 301 LWARSGLISSEEMAKELVELSKDKSKDNK 331

RESULT 3

US-09-815-242-12254
; Sequence 12254, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12254
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12254

Query Match 73.5%; Score 1242; DB 9; Length 337;
Best Local Similarity 72.0%; Pred. No. 2e-76;
Matches 236; Conservative 43; Mismatches 49; Indels 0; Gaps 0;

Qy 4 SVRGLKILSVIGLLPVLIIATAACGNSSSSSSKSSKDGVEIKHEEGTTKVPKPKRVV 63

```
Db 10 TMRGLKTSILGLI VALFLVAACGNTDSSKESSTKDTISVKDENGTVKVPKDAKRIW 69
Qy 64 LEYSFVDALVDVPGVGIADNKKRII KPLRDKIGKTVTSVGTQKNLEISKLPDL 123
Db 70 LEYSFADALALVDVPGVGIADNKKRII KPVREKIGNTVTSVGTQKNLEISKLPDL 129
Qy 124 IIAONNRHGIYKDLNKIAPTIELSKSFGDYNENIDAFKTI SKALGKEEGKRLAEHDK 183
Db 130 IIAUSSRHKGINKELNKIAPTIELSKSFGDGYKQNSFKTIAKALNKEKEGKRLAEHDK 189
Qy 184 KIEEYKEITMDKNOKVLPVAAKSGLLAHPNSVVGQFLSOLGPKFKEALSDDVTKGLSKY 243
Db 190 LIKKYKDEIKFDRNOKVLPVAVAKAGLLAHPNYSYVGQFLNELGFKNALSDDVTKGLSKY 249
Qy 244 LKGPVLOMNTETLSQVNPFRMIFMTNKASSNEPSLKELEKDPVWKKNVKNQORVDILDR 303
Db 250 LKGPVLOMNTETLSQVNPFRMIFMTNKASSNEPSLKELEKDPVWKKNVKNQORVDILDR 309
Qy 304 DLWARSGLISSEEMAKELVELSKKDSK 331
Db 310 DLWARSGLISSEEMAKELVELSKKDSK 337
```

```
RESULT 4
US-10-282-122A-43828
; Sequence 43828, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43828
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-43828
```

```
Query Match 72.9%; Score 1232; DB 15; Length 327;
Best Local Similarity 71.9%; Pred. No. 9.1e-76;
Matches 235; Conservative 42; Mismatches 50; Indels 0; Gaps 0;
Qy 5 VRGLKILSVITGLL FVLVIATAACGNSSNSKSSKSGDVEIKHEEGTTKVPKPKRVVVL 64
Db 1 MRGLKTSILGLI VALFLVAACGNTDSSKESSTKDTISVKDENGTVKVPKDAKRIW 60
Qy 65 BYSFVDALVDVPGVGIADNKKRII KPLRDKIGKTVTSVGTQKNLEISKLPDL 124
Db 61 EYSFADALALVDVPGVGIADNKKRII KPVREKIGNTVTSVGTQKNLEISKLPDL 120
Qy 125 IADNNRHGIYKDLNKIAPTIELSKSFGDYNENIDAFKTI SKALGKEEGKRLAEHDK 184
Db 121 IADSSRHKGINKELNKIAPTIELSKSFGDGYKQNSFKTIAKALNKEKEGKRLAEHDK 180
Qy 185 IEEYKEITMDKNOKVLPVAAKSGLLAHPNSVVGQFLSOLGPKFKEALSDDVTKGLSKYL 244
Db 181 INKYKDEIKFDRNOKVLPVAVAKAGLLAHPNYSYVGQFLNELGFKNALSDDVTKGLSKYL 240
Qy 245 KGPVLOMNTETLSQVNPFRMIFMTNKASSNEPSLKELEKDPVWKKNVKNQORVDILDR 304
Db 241 KGPVLOMNTETLSQVNPFRMIFMTNKASSNEPSLKELEKDPVWKKNVKNQORVDILDR 300
Qy 305 LWARSGLISSEEMAKELVELSKKDSK 331
Db 301 VWARSGLISSEEMAKELVELSKKDSK 327
RESULT 5
US-09-815-242-5459
; Sequence 5459, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5459
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5459
Query Match 69.9%; Score 1181; DB 9; Length 309;
Best Local Similarity 72.8%; Pred. No. 2.5e-72;
Matches 225; Conservative 38; Mismatches 46; Indels 0; Gaps 0;
```

Qy	16	LLFVLIATAACNNSSNSSKSSKGGVYEIKHEBGTTKVPKHPKRVVWLVESFVDALVAL	75
Db		: : : : : : : : : : : : : : : : : : : :	
1	1	LIVALFLVAACGNTDSSKSSSTKDTISVKDENGTVKVPKDAKRIVVLEVESFADALAAL	60
Qy	76	DVKPVGTADDNKNRIIKPLRDKIKGYTSVGRKOPNLEBESKLPDLDLIADNNRHKGIIY	135
Db		: : : : : : : : : : : : : : : : : : : :	
61	61	DVKPVGTADDGCKRIIKPVREKIGNYTSVGRKQPNLEBESKLPDLDLIADSSRHKGIN	120
Qy	136	KDLNKIAPTIELSKPFDGDNENIDAFKTSIALGKEEGSKRLEBHDKKIIBEYKKEITMD	195
Db		: : : : : : : : : : : : : : : : : : : :	
121	121	KELNKIAPTLSLKPFDGDKQNIINSFKTIAKALNKEGEGKRLAEHDKLIKKYKDEIKPD	180
Qy	196	KNQKVLPAVAKSGLLHAHPSSNYVGOFLSQLGFKFKEALSDDVTKGLSKYLKGPYLOMNTET	255
Db		: : : : : : : : : : : : : : : : : : : :	
181	181	RNQKVLPAVAKAGLLAHPNYSYVGQFLNELGFKNALSDDVTKGLSKYLKGPYLOLDETH	240
Qy	256	LSQVNPERRMFMTNKASSNPSLKELEKDPVWKKLNVAKNQRVDDILDRDLMAWRSRGLISS	315
Db		: : : : : : : : : : : : : : : : : : : :	
241	241	LADLNPERMIITMDNAKDSAEFKQLQEDPTWKKLNVAKNRNRVDIVDRDVMWRSRGLISS	300
Qy	316	REMAKELVE	324
Db		: : : : : : : : : : : : : : : : : : : :	
301	301	REMAKELVE	309

RESULT 6

```

US-09-815-242-5249
Sequence 5249, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yanamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Esse
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5249
LENGTH: 319
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5249

```

```

Qy 66 YSFVDALVALDVKPVGIADDDKKRRIKPLRDKTKGKTVSGTRKQPNLBEISKLKPDLLI 125
Db 59 YSFADYLAALDDMKVGIADDDGSSKNITKSVDRDKVAYESVGSRPQPNMEVTSKLKPDLLI 118
Qy 126 ADNRRHGKIVYKDLNKNIAPTIELKSPDGDYNNENIDAFKTIKALGKEEKGKRLKEHDKKI 185
Db 119 ADVSRHKIKSELSKIAPTIMLVSGTDYNNANIADFKTVAKAVGKEKGKRLKEHDKIL 178
Qy 186 EYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQGFLSQLGFKALSDDVTKGLSKYLK 245
Db 179 AEIRKKEIQSLTKTAFAGISRAGWFINNETFMQFLIKMGIQPEVTKDKTTHVGERKG 238
Qy 246 GPYLOMNTETLSQVNPBRMFTW-N-KASSNPSPLKELEKDP-VMKKLNAAVKNQRVDILDR 303
Db 239 GPYTYLANNEELANTNPVMILATDGTDKNRTKF-----IDPAVMSKSLKAVKONKQVYDVDR 294
Qy 304 DLWARSRLGISSEEMAKELVELSKK 328
Db 295 NKWLKSRGIATSESAEDLEKIAEK 319

RESULT 7
US-09-815-242-12122
; Sequence 12122, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ IDS NOS: 14110
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 12122
; LENGTH: 319
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-12122

```

RESULT 7

```

US-09-815-242-12122
; Sequence 12122, Application US/09815242
; Patent No. US20020061569A1
GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: EUTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12122
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12122

```

```

Query Match      43.0%; Score 726; DB 9; Length 319;
Best Local Similarity 47.1%; Pred. No. 2.2e-41;
Matches 153; Conservative 61; Mismatches 99; Indels 12; Gaps 5;

Qy 6 RGLKILSVIGLFLVLIATAACGNSSNSSKSSKDCGVEIKHEGTTKVPKPKRVRVVLE 65
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5 RNIKLVVFMILIFV-VAVAGGGQKDT-----ERKTEMTTIKDELGETIKKPKRVRVVLE 58
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Qy	6	RGLKILSVIGGLFVLIIATAAGNNSSNSKESKOGVEIKHEEGITTPVKPKHGVVLE	65
Db	3	RNIIVKLVFVLFV-VAVAGCGQKDT-----EKEEMTTIKDELGTSEKIKNPKRVVLE	56
Qy	66	YSPVDALVALVDKVPVGVIADGNKNRIIKPLRDKIKGYTSVGTGRQPNLEESIKLKPDLII	125
Db	57	YSFADYLAALDMKPVGIADGGSKNIITPSVRDKVGYVESYGRSQPNNEVISKLKPOLII	116

QY	126	ADNNRHGIYKDLNKIAPTIELKSPDGYNENIDAFKTIISKALGKEEGKRLBEHDKKI	185
Db	117	ADVSRHKIKSELSKIAPTIMLVSGTDYNNANIDAFKIVAKAVGKEGKRLBEHDKKIL	176
QY	186	BEYKKEITMDNKNQKVLPAVAKSGLLAHPNSYVGQFLSQGLFKALSDDVTYKGLSKYLK	245
Db	177	AEIRKKIEQSTLKTAFAFGISRAGMFINNEDTFMGQFLIKMGIOPEVTKDTHVGERKG	236
QY	246	GPYLOMNTETLSQVNPFRMFINN-KASSNEPSLKELEKDP-VWKKLNAVKNQRYDILDR	303
Db	237	GPYIYNNEELANINPKVMILATDGTDKNRTKF-----IDPAVWKSLLKAVKDNKYVDVDR	292
QY	304	DLWARSRLGISSEEMAKELVELSKK	328
Db	293	NKWLKSRGLIASMAEDLEKIAEK	317
RESULT 8			
US-10-282-122A-44133			
; Sequence 44133, Application US/10282122A			
; Publication No. US20040029129A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Liangseu			
; APPLICANT: Zamudio, Carlos			
; APPLICANT: Malone, Cheryl			
; APPLICANT: Haselbeck, Robert			
; APPLICANT: Ohlsen, Kari			
; APPLICANT: Zyskind, Judith			
; APPLICANT: Wall, Daniel			
; APPLICANT: Trawick, John			
; APPLICANT: Grant			
; APPLICANT: Yamamoto, Robert			
; APPLICANT: Forsyth, R.			
; APPLICANT: Xu, H.			
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms			
; FILE REFERENCE: ELITRA.034A			
; CURRENT APPLICATION NUMBER: US/10/282,122A			
; CURRENT FILING DATE: 2003-02-20			
; PRIOR APPLICATION NUMBER: 60/191,078			
; PRIOR FILING DATE: 2000-03-21			
; PRIOR APPLICATION NUMBER: 60/206,848			
; PRIOR FILING DATE: 2000-05-23			
; PRIOR APPLICATION NUMBER: 60/207,727			
; PRIOR FILING DATE: 2000-05-26			
; PRIOR APPLICATION NUMBER: 60/230,335			
; PRIOR FILING DATE: 2000-09-06			
; PRIOR APPLICATION NUMBER: 60/230,347			
; PRIOR FILING DATE: 2000-09-09			
; PRIOR APPLICATION NUMBER: 60/242,578			
; PRIOR FILING DATE: 2000-10-23			
; PRIOR APPLICATION NUMBER: 60/253,625			
; PRIOR FILING DATE: 2000-11-27			
; PRIOR APPLICATION NUMBER: 60/257,931			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: 60/267,636			
; PRIOR FILING DATE: 2001-02-09			
; PRIOR APPLICATION NUMBER: 60/269,308			
; PRIOR FILING DATE: 2001-02-16			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 78614			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 44133			
; LENGTH: 319			
; TYPE: PRT			
; ORGANISM: Staphylococcus aureus			
US-10-282-122A-44133			
Query Match 42.9%; Score 724.5; DB 15; Length 319;			
Best Local Similarity 46.9%; Pred. No. 2.8e-41;			
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;			
QY	13	VIGLLFVLATACGNSSNSSKSSKDGVEIKHEEGTKVPHKPRVVVLEYSFVDAL	72

```

US-10-282-122A-71291
; Sequence 71291, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71291
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71291

Query Match      38.1%; Score 644; DB 15; Length 175;
Best Local Similarity 70.9%; Pred. No. 4e-36;
Matches 124; Conservative 23; Mismatches 24; Indels 4; Gaps 1;

Qy      5 VRGKILSVIGLLFVLIIATAACGNSSNS----SKESSKDGVEIKHEEGTTKVPKPKR 60
Db      1 MKGFAGIVALLFALVLVITACGNVNNCGSDSGNKKSSKDSIKIKHELGTTKVPKDAKR 60

Qy      61 VVVLSEYSPDALVDVKPVGIADNKKRIIKPLRDKIGKTVSGTRKQPNLEISKJK 120
Db      61 VVALEFSPDALAALNVKPVGVADNKKRIIKPLKEKIGDKYKSVGARKQPNLEISKJK 120

Qy      121 PDLIADNNRHGIYKDLNKIAPTIELKSPGPDYNNENIDAPKTIKALGKEEGK 175
Db      121 PDLIADNNRHGIYKELSKIAPTIELKSPGPDYNDNDIDAPKTIKALNKDDVQG 175

RESULT 11
US-10-282-122A-66890
; Sequence 66890, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos

```

```

; APPLICANT: Malone, Cheryl
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66890
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-66890

Query Match      30.6%; Score 516.5; DB 15; Length 298;
Best Local Similarity 38.1%; Pred. No. 3.6e-27;
Matches 111; Conservative 65; Mismatches 98; Indels 17; Gaps 4;

Qy      43 VEIKHEEGTTKVPKPKRVVVLSEYSPDALVDVKPVGIADNKKRIIKPLRDKIGKY 102
Db      23 VTVKQKGEFTLDSVPKRVVALEYSYVDALAQIGVSPGVADDDNKTILQKVRDKVQPW 82

Qy      103 TSVGTRKQPNLEISKJKLPDLIIADNNRHGIYKDLNKIAPTIELKSPGPDYNNENIDAFK 162
Db      83 ESVGTRSQPSLEAISALKPDLIIADNNRHSAYVELKXIAPTIVFNSRSHENYQENLETAQ 142

Qy      163 TISKALGKEEGKRLSEHDKKIEEYKEITMDKNQVLPVAAKSGLLAHPNSYSYVQOF 222
Db      143 KIGOLLGKREMQARIATKHQODIADIAK--TLPGKKAIIIGVSRQTNLNSRSGYAGGL 200

Qy      223 LSQLGFK--EALSDDVTGKLSKYLKGPVLOQNTETLSQVNPFRMIFMTNKASSRPSLKE 280
Db      201 VEVLYGQMPKARADN-----QPNASVGLGQVAAEKPDMLILHYR---DESIARK 247

Qy      281 LEKDPVWKULNAVQNRQVDILDRDLWARSGLISSEMAKELVELSKDSK 331
Db      248 WENEALWKIIPAVKNGQVILANDNLWARAGIDAAEVMAKEVQDFVTKSAK 298

RESULT 12
US-10-282-122A-56054
; Sequence 56054, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

```



```

; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56054
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-56054

Query Match      27.2%; Score 460; DB 15; Length 300;
Best Local Similarity 35.8%; Pred. No. 2.5e-23;
Matches 101; Conservative 60; Mismatches 105; Indels 16; Gaps 4;

QY 45 IKHEEGTTKPKPKRVVLEYSFVDALVDPKVGADDDNNKRIIKPLRDKIGKYS 104
DB 24 VQDEHGTFTLTKTPQRIIVVLEISFADALAAVDVPIGADDDNNKRIILPEVRAHLKPMQS 83

QY 105 VGTROPNLEESIKLPDLIIADNNRHKGIIYKDLNKIAPTIELKSPDGDYNNIDAFKTI 164
DB 84 VGTQAQPSLEATAALPDLIIADSSRHAGIYALQOIAPVLLJKSRNETYAENLHSAAT 143

QY 165 SKALGKEEGKKRLKEHDKKIEEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQGFUS 224
DB 144 GEVVGKKREMQARLEQHKERMAQWSQ--LPTGTRVAFGTSRQLFNLTQETWTGSLVLA 201

QY 225 QLGFKALSDDDVTGKLSKYLKGPYL-QMNTETLSQVNPBRMFTMTNKASSNEPSLKELEK 283
DB 202 SLGL-----NVPAAAGASMPISGLEQLLAVNPALVLAHVYR---EESIVKRWQ 248

QY 284 DPVWKKLNAVKNQVRDILDRDLWARSGLISSEEMAKELVEL 325
DB 249 DPLWQILTAQAQKQVAVSDSNARWGRGIFAAERIAADTVKI 290

RESULT 13
US-09-815-242-10469
; Sequence 10469, Application US/09815242

; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10469
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10469

Query Match      27.2%; Score 460; DB 9; Length 302;
Best Local Similarity 35.5%; Pred. No. 2.5e-23;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;

QY 45 IKHEEGTTKPKPKRVVLEYSFVDALVDPKVGADDDNNKRIIKPLRDKIGKYS 104
DB 26 VQDEHGTFTLTKTPQRIIVVLEISFADALAAVDVPIGADDDNNKRIILPEVRAHLKPMQS 85

QY 105 VGTROPNLEESIKLPDLIIADNNRHKGIIYKDLNKIAPTIELKSPDGDYNNIDAFKTI 164
DB 86 VGTQAQPSLEATAALPDLIIADSSRHAGIYALQOIAPVLLJKSRNETYAENLHSAAT 145

QY 165 SKALGKEEGKKRLKEHDKKIEEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQGFUS 224
DB 146 GEVVGKKREMQARLEQHKERMAQWSQ--LPKGTRVAFGTSRQLFNLTQETWTGSLVLA 203

QY 225 QLGFKALSDDDVTGKLSKYLKGPYL-QMNTETLSQVNPBRMFTMTNKASSNEPSLKELEK 283
DB 204 SLGL-----NVPAAAGASMPISGLEQLLAVNPALVLAHVYR---EESIVKRWQ 250

QY 284 DPVWKKLNAVKNQVRDILDRDLWARSGLISSEEMAKELVEL 325
DB 251 DPLWQILTAQAQKQVAVSDSNARWGRGIFAAERIAADTVKI 292

RESULT 14
US-10-282-122A-56805
; Sequence 56805, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56805
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-56805

Query Match      27.2%; Score 460; DB 15; Length 302;
Best Local Similarity 35.5%; Pred. No. 2.5e-23;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;

Qy 45 IKHEGTTKPKPKRVVVLVSFVDALVDKPVGIADNKKNRITKPLDKIGKYS 104
Db 26 VQDEHGTFTLTKTPQRIYVLSLSPADALAAVDVIPGIGIADNDKAKRILLPEVRAHLKPWQS 85
Qy 105 VGTRKQPNLEISKLPDLITADNNRHGKIYKDLNKNKIAPTIELKSFQDGYNENIDAFKTI 164
Db 86 VGTRAQPSLEIAALKPDLIIADSSRHAGVIALQOIAPVLLKSRNETYAENLQSAAIL 145
Qy 165 SKALGKEEGKRLEEDKIEYEKKEITMDKNQKVLPAVAASKGLLAHPSNSVYGFLS 224
Db 146 GEMVQKKREMQARLEQHKERMAQWASQ--LPKGRVAFGTSREQQFNLTQETWTGSLA 203
Qy 225 QLGFKALSDVTGSLKYLKGPYL-QMNTETLSQVNPFRMFTNKASSNEPSPLEK 283
Db 204 SLGL-----NVPAAAGASMPSTGLQLAVNPALLVAHYR---EESIVKRWQQ 250
Qy 284 DPVKKLVNAKNQVRIIDRLDLWARSGLISSEEMAKSELVEL 325
Db 251 DPLWQLMTAAQKQVQVAVSDSNWARMRGIPAAERTADTVKI 292

RESULT 15
US-10-282-122A-69420
; Sequence 69420, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haseelbeck, Robert
```

```
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Yamamoto, Robert
; APPLICANT: Carr, Grant
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69420
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69420

Query Match      26.9%; Score 454; DB 15; Length 297;
Best Local Similarity 34.3%; Pred. No. 6.3e-23;
Matches 109; Conservative 60; Mismatches 109; Indels 40; Gaps 6;

Qy 17 LFVLIATAACGNSSNSSKSSKSSKDGVEIKHEGTTKV--PKHPKRVVVLVSFVDALVA 74
Db 5 LLTLTLAGAA-----QAAPIDIDDDGQHKVHLPTPKRVVVLVSFVDGLAS 49
Qy 75 LDVKEPVGIADNKKNRITKPLDKIGKYSVTRKQPNLEISKLPDLITADNNRHKGI 134
Db 50 VGVTFVGAADPGDASRVLPKVRKAVGEMQSVGLRSQPNIRVIAKLPDLIIADLGRHQA 109
Qy 135 YKDLNKIAPTIELKSFQDGYNENIDAPKTIISKALGKEEGKRLEEDKIEYEKKEITM 194
Db 110 YNDLASLAPTMLPSRGEDYQGLSKSAGLIGALGKGPQMAAENRQHLKTVABQIPA 169
Qy 195 DKNQKVLPAVAASKGLLAHPSNSVYGFQPSQLGFKALSDVTGSLKYLKGPYLMNTE 254
Db 170 DSN--VLPFGVAREDSFVHGPHSYAGSVLQIG-----LQVPEVRNNA 211
Qy 255 TLSQVNPFRMFTNK-----ASSNEPSLKEK-EXDPVKKLVNAKNQVRIIDRLDLWARS 309
Db 212 PTEFVSLEQLLALDPNMLLVGHYRRPSIVDTWSKQPLMQVLGAVRNKQVAVEDGDSWARN 271
Qy 310 RGLISSEEMAKSELVELSK 327
Db 272 RGIWASEQIADDAVLK 289

RESULT 16
US-08-781-986A-5193
; Sequence 5193, Application US/08781986A
```



```

/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/329,624
/ FILING DATE: 27-Dec-2002
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/956,171
/ FILING DATE: October 20, 1997
/ APPLICATION NUMBER: 60/009,861
/ FILING DATE: January 5, 1996
/ APPLICATION NUMBER: 08/781,986
/ FILING DATE: January 3, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mark J. Hyman
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB248P1D1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (240) 314-1224
/ TELEFAX: (301) 309-8439
/ INFORMATION FOR SEQ ID NO: 5193:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 330 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5193:
US-10-329-624-5193

Query Match      24.0%; Score 405; DB 15; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.5e-19;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY  8 LKILSVIGLLFVLITATAACGNSSNSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYS 67
DB  5 IKML-VVTLAFLV-LAGCSGNSNKQSSDNKDKEKTTTSIKHAMGTTTEIKGPKRVVTLVYQG 62
QY  68 FVDALVALDVKPGVIADDNKKNIIRKPLRDKIGKTVSGTRKQPNLEISKLKPDLIAD 127
DB  63 ATDVAVSLGVKPGVAVESWTQKPKFEYIKNDLKDTKIVGQBPAPNLEISKLKPDLIAS 122
QY  128 NNRHKGIYKDLNKIAPTIELKSPGDYNNENIDAFKTIISKALGKEEGKRLKEEHDKKIEE 187
DB  123 KVRNEKVVYDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEKEAEEDLLKKYDDKVA 178
QY  188 YKKEITMDKNQKVLPAVAKSGLLAHPSN-----SYVGQFLSGLGPKKEALSDDVTKGL 240
DB  179 FQK-----DAKAKYKDAWPLKASVNVFRADHTRIYAGGYAGETLNDLGPKR-----NKDL 228
QY  241 SKYLAG--PYLQMT--ETLSQVNPFRMFTMTNKASSNEPSL-----KELEKDPVWKLN 292
DB  229 QKQVDNGKDIILQTSKESIPLMNADHIFVVKSDPNKADALVKKTESEWTSSEWTSSEWKNLDA 288
QY  293 VKNQRV-DILDRDLWARSGRLISSEMAKELVELSKDKSKDNK 335
DB  289 VKNQVSDDLDEITWNLAGGYKSSKLIDDLVE--KLNIEKQSK 330

RESULT 19
US-10-470-048B-346
; Sequence 346, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; FILE REFERENCE: ANTIGENS TO A SPECIFIC PATHOGEN
; CURRENT APPLICATION NUMBER: US/10/470,048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603

Query Match      24.0%; Score 405; DB 15; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.5e-19;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY  8 LKILSVIGLLFVLITATAACGNSSNSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYS 67
DB  5 IKML-VVTLAFLV-LAGCSGNSNKQSSDNKDKEKTTTSIKHAMGTTTEIKGPKRVVTLVYQG 62
QY  68 FVDALVALDVKPGVIADDNKKNIIRKPLRDKIGKTVSGTRKQPNLEISKLKPDLIAD 127
DB  63 ATDVAVSLGVKPGVAVESWTQKPKFEYIKNDLKDTKIVGQBPAPNLEISKLKPDLIAS 122
QY  128 NNRHKGIYKDLNKIAPTIELKSPGDYNNENIDAFKTIISKALGKEEGKRLKEEHDKKIEE 187
DB  123 KVRNEKVVYDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEKEAEEDLLKKYDDKVA 178
QY  188 YKKEITMDKNQKVLPAVAKSGLLAHPSN-----SYVGQFLSGLGPKKEALSDDVTKGL 240
DB  179 FQK-----DAKAKYKDAWPLKASVNVFRADHTRIYAGGYAGETLNDLGPKR-----NKDL 228
QY  241 SKYLAG--PYLQMT--ETLSQVNPFRMFTMTNKASSNEPSL-----KELEKDPVWKLN 292
DB  229 QKQVDNGKDIILQTSKESIPLMNADHIFVVKSDPNKADALVKKTESEWTSSEWTSSEWKNLDA 288
QY  293 VKNQRV-DILDRDLWARSGRLISSEMAKELVELSKDKSKDNK 335
DB  289 VKNQVSDDLDEITWNLAGGYKSSKLIDDLVE--KLNIEKQSK 330

RESULT 19
US-10-470-048B-346
; Sequence 346, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; FILE REFERENCE: ANTIGENS TO A SPECIFIC PATHOGEN
; CURRENT APPLICATION NUMBER: US/10/470,048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 346
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-470-048B-346

Query Match      24.0%; Score 405; DB 17; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.5e-19;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY  8 LKILSVIGLLFVLITATAACGNSSNSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYS 67
DB  5 IKML-VVTLAFLV-LAGCSGNSNKQSSDNKDKEKTTTSIKHAMGTTTEIKGPKRVVTLVYQG 62
QY  68 FVDALVALDVKPGVIADDNKKNIIRKPLRDKIGKTVSGTRKQPNLEISKLKPDLIAD 127
DB  63 ATDVAVSLGVKPGVAVESWTQKPKFEYIKNDLKDTKIVGQBPAPNLEISKLKPDLIAS 122
QY  128 NNRHKGIYKDLNKIAPTIELKSPGDYNNENIDAFKTIISKALGKEEGKRLKEEHDKKIEE 187
DB  123 KVRNEKVVYDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEKEAEEDLLKKYDDKVA 178
QY  188 YKKEITMDKNQKVLPAVAKSGLLAHPSN-----SYVGQFLSGLGPKKEALSDDVTKGL 240
DB  179 FQK-----DAKAKYKDAWPLKASVNVFRADHTRIYAGGYAGETLNDLGPKR-----NKDL 228
QY  241 SKYLAG--PYLQMT--ETLSQVNPFRMFTMTNKASSNEPSL-----KELEKDPVWKLN 292
DB  229 QKQVDNGKDIILQTSKESIPLMNADHIFVVKSDPNKADALVKKTESEWTSSEWTSSEWKNLDA 288
QY  293 VKNQRV-DILDRDLWARSGRLISSEMAKELVELSKDKSKDNK 335
DB  289 VKNQVSDDLDEITWNLAGGYKSSKLIDDLVE--KLNIEKQSK 330

RESULT 20
US-10-967-189-16
; Sequence 16, Application US/10967189
; Publication No. US20050053995A1
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USD2
; CURRENT APPLICATION NUMBER: US/10/967,189
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: US 10/278,946
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-967-189-16

Query Match      24.0%; Score 405; DB 17; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.5e-19;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY  8 LKILSVIGLLFVLITATAACGNSSNSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYS 67
DB  5 IKML-VVTLAFLV-LAGCSGNSNKQSSDNKDKEKTTTSIKHAMGTTTEIKGPKRVVTLVYQG 62

```

QY 68 FVDALVALDVKPGVGIADNDKNNRIIKPLDKIGKYSVCTRKQPNLEISLKPDLIIAD 127
 DB 63 ATDVAVSLGVKPGVAVESWTQPKFEYIKNDUKOTKIVQBPAPNLEISLKPDLIIAS 122
 QY 128 NNRHGIYKDLNKIAPTIELKSFDDGYNENIDAFKTIKSKALGKEEGKRLKEHDKKIEE 187
 DB 123 KYRNEKVDOLSKIAPTV---STDIVPKFK-DTTLKMGKALGKEAEADLLKKYDDKVA 178
 QY 188 YKKEITMDKQKVLPAVAKSGLLAHPSN-----SYGQFLSOLGPKKEALSDDVTYGL 240
 DB 179 FQK----DAKAKYKDAWPLKASVFNFRADHTRIYAGGYAGEIINDLGFKR-----NKDL 228
 QY 241 SKYLGK--PYLQWNT--ETLSQVNPERRMTNKAASNEPSL-----KELEKDPVWKCLNA 292
 DB 229 QKQVNGKDIILQTSKESILPMNADHIFVVKSDPNKADALVKKTESEWTSSTSEKWNLDA 288
 QY 293 VKNQV--DILDRDLWARSGLISSEEMAKELVELSKOSKDK 335
 DB 289 VKNQVSDDLDEITWNLAGYSKSLKLIIDLYE--KLNIKQSK 330

RESULT 21
 US-10-282-122A-45599
 ; Sequence 45599, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangau
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 45599
 ; LENGTH: 324
 ; TYPE: PR1
 ; ORGANISM: Bacillus anthracis
 US-10-282-122A-45599

Query Match 23.1%; Score 391; DB 15; Length 324;

Best Local Similarity 34.4%; Pred. No. 1.4e-18;
 Matches 116; Conservative 59; Mismatches 122; Indels 40; Gaps 13;
 QY 9 KILSVIGLLFVLIIATAACGNSSSSSKESKD---GVEIKHEGTTKVPKPKRVVVL 64
 DB 3 KILSI---FIVWFLFVAGCCQKEEKKETKADKNQAITIKHAEGETKLDKPAKVVVL 58
 QY 65 EYSFVDALVALDVKPGVGIADNDKNNRII---KPLRDKIGKYSVGTTRKQPNLEISLKL 120
 DB 59 EMWYSEDLLAGVQVPGVGMADIKNNKWNVTNTPKSKDVV---DVGTRQPNLEISLKL 114
 QY 121 PDLIIADNNRUKGIYKDLNKIAPTIEL---KSFDDGYNENIDAFKTIKSKALGKEEGKRL 177
 DB 115 PDLIIITASFRGKAIKNELEQIAPTVMFDPSTNNNDHFAEMTETPKQIAKAVKSEEGKKV 174
 QY 178 LEEDHKKIEEYKKEI---TWDKQKVLPAVAKS---GLLAHPSNSYVGQFLSOLGFK 229
 DB 175 LADMDKAFADAKAKIEKADLKDKKNIAMAQAATAKVNPTFRILT--DNSLAQVTKLGG-- 230
 QY 230 EALSDDVTGKLSKYLKGPYLMNTTETLSQVNPERRMTNKAASNEPSLKELEKDPVWK 288
 DB 231 --LNTTPEAGKSE--PDGFKQTTVESLQSVQDSNFIYIVADEDNIFDTQLK---GNPAWE 283
 QY 289 KLNVAQNORVDILDRDLWARSGLISSEEMAKELVEL 325
 DB 284 ELKFKKENKMYKLGDTWIFG-GPESATSLATQVADV 319

RESULT 22
 US-09-738-626-3875
 ; Sequence 3875, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 3875
 ; LENGTH: 315
 ; TYPE: PR1
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3875

Query Match 19.6%; Score 331; DB 9; Length 315;
 Best Local Similarity 28.6%; Pred. No. 1.6e-14;
 Matches 90; Conservative 65; Mismatches 140; Indels 20; Gaps 7;
 QY 13 VIGLLFVLIIATAACGNSSSSNS-SKESKQGVKEIKHEGTTKVPKPKRVVVLSEYFVDA 71
 DB 15 VAPLLGVALLSSCSSTSSDESIGQPEVASTGYGVEHAMGTETPTPTPTVWIDSPHLD 74
 QY 72 LVALDVKPGVGIADNDKNNRIIKPLRDKIGKYSVGTTRKQPNLEISLKLKPDLLIADNNRH 131
 DB 75 LLALGITPVGATESGSENGFPAYLADELKDTESVGLTSEPNLEKIAALPDLLIAGKVRH 134


```
Query Match      18.5%; Score 313; DB 15; Length 314;
Best Local Similarity 29.9%; Pred. No. 2.6e-13;
Matches 99; Conservative 65; Mismatches 135; Indels 32; Gaps 14;

QY 9 KILSVIGLLFVLIATAACGNSSNSSKESKDGVEIKHEEGTTKVPKHPKRVVLEYSF 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 KLFILFTMLVLSVIGSSSQESKAKEQPKTKV-VKHAKGEATIPVNPRIIVDLSGS- 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 VDALVALDVKPGIADNNKRIIKPLDKIKGYTSVGRKQP--NLREISKLPDLIIA 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 TEELLLGHKPGVGT-A-NYKDKIQKHLTEKLDGVKAVGWYAPKVDLEAVTALKPDLIL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 DNNRHGYYKOLNIAPIELKSPDQYNENIDAFKTIKSKALGKEEGKGLLEHDKKIE 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 -NNRQLKIYDQLEKVAPTVVLNLEDMR---GKFEVGLFDEKKADKIADYDKKAD 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 EYKKEITWDKQKVLPAVAAGSGLLAHPSNSV-GQFLSQGLGFKALSDDV-----TKGLS 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 SLSKKIKKTKDSDSFMFAVT-----PQNFVYGSF---GYGDIIFNDLKLPAKGTG 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KYLKGPIQMTTETLSQVNPFRMFTNKNASSNEPSLKELEKDPYWKLNNAVKNQV 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 --LKQTAQVLSGLVAFQPDQMFIV-NFGGEADKVYEDYKNSAVKONKAVNNHVEV 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 DRDLW-ARSRGLISS-----EEMAKELVELSK 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 SNEVNTKAFNPIGKMLDEIAKELILAKNK 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 25

```
US-10-282-122A-45385
; Sequence 45385, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 45385
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-45385

Query Match      17.8%; Score 301; DB 15; Length 314;
Best Local Similarity 27.3%; Pred. No. 1.7e-12;
Matches 93; Conservative 66; Mismatches 106; Indels 76; Gaps 16;

QY 16 LLFVLIIAT----AACGNSSSNS--SKESKDGVEIKHEEGTTKVPKHPKRVVLEYSFVD 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 LAWLVLVTSVLFACSNKKEKADAKAETERTVQHAKEIKIPANPKKIADLSGS-TE 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 ALVALDVKPGIADNN-----KKNRIIKPL-----RDKIGKYTSVGRKQPNLEIS 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 ELLIFGKMKPIITANTSQEKIDAHEIKELKQKVPVGSAGWDKI-----NIEAVA 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 KLPDLIIADNNRHGYYKOLNIAPIELKSPDQYNENIDAFKTIKSKALGKEEGK 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 AAKPDLILV--NNRQEKIYDQLSKIAPTVMKLT---PLDQWRPKFEVGGQIFGK 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 LEEDHKKEIEYKKEITWDKQKVLPAVAAGS-----LLAHPNSVYVGQFLSQGLFKEA 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 FKQYD-----EKASKLHDKIVAKTGDAKFMKMAAYP--NAP--RVYGDYGYGSV 216
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 232 LSDDV-----TKGLSKYLKGPYLQMTTETLSQVNPFRMFTNKNASSNEPSLKELEKOPVW 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 IFNDLKLPAVKGTP--TDKPLVQVKEALIDYNPDYLFVFTTGDGDSQR--LKEFQ 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 KCLNAVKNQVRVDILDRDLWARSRLGISSEEMAKELVELSKK 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 KMNNAVKNHV-----FTIKNEELNKGYPFLGKE 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 26

```
US-10-282-122A-46193
; Sequence 46193, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
```


Query Match 17.5%; Score 296; DB 15; Length 342;
Best Local Similarity 29.2%; Pred. No. 4.2e-12;
Matches 90; Conservative 66; Mismatches 118; Indels 34; Gaps 11;
QY 9 KILSVIGLLFVLIIATAACGNNSSSSKSSKGGVEIKHEEGTTKVPKPKRVVWVLEYSF 68
DB 27 KSLSPFLFSLVLAGCDNAQDTSTQSTQTEKQTLTIEHAQQTTEIPAHPOKVVMMNET 86
QY 69 VDALVALDV--KPGVIADNKKNRRIKPLRDKTGKYSVGTGRKOPNLEESKLPDII 122
DB 87 LDIVDALGVVVGLPQTNVH-----LPKLEKYNENYINEGGLFPNTEKLSSTAPD 140
QY 123 LIADNRHGIYKDLNKIAPTIELKSPGDYNNENIDAF-----KTISKALGKEEGKRL 178
DB 141 LLT--GSRADAYAKLSEIAPSI---SMDIDSTRIDSITERTTTLGQIFGKEQAKLL 196
QY 179 EEHDKKIEYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVQFLSQLGFKKALSDDVTK 238
DB 197 ADPNSKIDTVKAK--TPDAG--KAMVVLVSGGKISAYGPGSRFGFIYDVLGFPAYTFD--SP 253
QY 239 GLSKYLKGPYLOMNTETLSQVNPFRMFM--TNKASSNEPSLKELEKDPVWKKLNAVX 294
DB 254 GSHGNI-----VNSELLKLPDMFMFVIDRDAATGREDSDPA--KQVLDNALVRKVNWN 306
QY 295 NORVDILD 302
DB 307 KQIIYILD 314

RESULT 29
US-09-071-035-38
; Sequence 38, Application US/09071035
; Publication No. US20020045737A1
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-38

Query Match 16.7%; Score 283; DB 9; Length 317;
Best Local Similarity 30.4%; Pred. No. 2.9e-11;
Matches 101; Conservative 59; Mismatches 130; Indels 42; Gaps 13;
QY 11 LSVIGLLFVLIIATAACGNNSSSSKSSKGGVEIKHEEGTTKVPKPKRVVWVLEYS 67
DB 11 VSMGLML-----SACQTNKKTADSATETVKTGNTGQLTVPKPKRVVFDNG 66
QY 68 FVDALVALDV--KPGVIADNKKNRRIKPLRDKTGKYSVGTGRKOPNLEESKLPDII 125
DB 67 SLDTMDALGVGDVRVVGAPTKN-----IPAYLKTKYKVESAGGIKEPDKLEKINQLKPDII 121
QY 126 ADNRHGIYKDLNKIAPTIELKSPGDYNNENIDAFKTISKALGKEEGKRL 181
DB 122 I--SGRQDYQQLKAIAPTIIYLAVDKPNWASTKQNTIETLGTI---FDKEEVAKEKITGL 177
QY 182 DKIEEYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVQFLSQLGFKKALSDD---VT 237
DB 178 EKEIADVKQAEASAN--NALVVLNEGQLSAYGSGRFLIHDTFGKAA--DDKIEAST 234
QY 238 KGLSKYLKGPYLOMNTETLSQVNPFRMFM--TNKASSNEPSLKELEKDPVWKKLNAVXQ 296
DB 235 HGQS-----VSIEYVLEKNGFILFVVDRTKAIGGDDSKDNVAANELIQTDAKND 285
QY 297 RVDILDRDLWARSGLISSEEMAKELVELSKK 328
DB 286 KVMILQPDVWVYLSGGGLSEMH---MIEDVKK 314
RESULT 30
US-10-206-576-38
; Sequence 38, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein


```
; ORGANISM: Enterococcus faecalis
US-09-815-242-10905

Query Match      16.5%; Score 279; DB 9; Length 317;
Best Local Similarity 30.1%; Pred. No. 5.4e-11;
Matches 100; Conservative 60; Mismatches 130; Indels 42; Gaps 13;

QY 11 LSVIGLLFVLIIATAACGNN---SSNSKSKSSKOGVEIKHEEGTTKPKPKRVVVVLEYS 67
Db 11 VSMGLMLL-----SACQTKTKTADATTTAKTEVTVDNGQLTVPKNQKVVVFDNG 66
QY 68 FVDALVALDV--KPGVIGIADDNKKRIIKPLRDKIKGYTSVGRKQPNLEESKLPDII 125
Db 67 SLDTMDALGVGRVVVGPATKN-----IPAYLKKYQKVESAGGIKEPDKINQLKPDII 121
QY 136 ADNRRHKGYIKDLNKIAPTIEL-----KSPDGDYNNENIDAFKTIISKALGKEEGKKRLSEH 181
Db 122 I-SGRQDQYQEQLEKAIAPTIIYLAVDAKNPWASTKQNIETLGTI-----FKKEEVAKEKITGL 177
QY 182 DKKIEYKKEITMDKNQKVLPAVAASKGLLAHPSNSYVQFLSQGLGFEALSDS----VT 237
Db 178 EKXIADVKKQASAN-NALVVLVNEGQLSAVKGSRFGLIHDTGFKAA--DDKIEAST 234
QY 238 KGLSKYLKGPYLOMNTETLSQVNPFRMFM-TNKASSNEPSLKELEKDPVWKKLVAVKQ 296
Db 235 HQQS-----VSVEYVLEKNPGILFVVDRTKAIGGDDSKDNVAANELIKOTDAGKND 285
QY 297 RVDILDRDLWARSGLISSEEMAKELVELSKK 328
Db 286 KVIMLQPDVWYLSGGGLSEMHIL---MIEDVKK 314

RESULT 33
US-10-282-122A-46179
; Sequence 46179, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekand, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46179
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-46179

Query Match      16.2%; Score 274.5; DB 15; Length 315;
Best Local Similarity 27.4%; Pred. No. 1.1e-10;
Matches 92; Conservative 70; Mismatches 129; Indels 45; Gaps 13;

QY 12 SVIGLLFVLIIATAACGNNSSNSKSKSSKOGVEIKHEEG-TTKVPKPKRVVVVLEYSFVD 70
Db 10 AILSIFILLIGCSAKGDEKASATKTEGKEKIEVTDLSGRKVTFDKVPESFATLSMGDMN 69
QY 71 ALVALDVKPGIADDNKKRIIKPLRDKIKGYTSVGRKQPNLEESKLPDIIADNRR 130
Db 70 IHALGGKIVGRPD-----AKITLP--EDIKKIQVIGNAHQNFQIASKLPDVLIIANGF 123
QY 131 HKGIYKDLNKIAPTIE-----LKSPDG--DYNENIDAFKTIISKALGKEEGKKRLSEH 181
Db 124 QKNI-----PIVEGGQTKVMISSANSVQDIQKNIELYGTIMK---KEDKANELNQKI 172
QY 182 DKKIEYKKEITMDKNQKVLPAVAASKGLLAHPSNSYVQFLSQGLGFEALSDSDDVTKGLS 241
Db 173 NVQMKYKYEK---SDVKALLVYGAPGYLAALPTSLSGDILEKTGGKNIAD--FPEMK 226
QY 242 KYLKGYPYLOMNTETLSQVNPFRMFM-TNKASSNEPSLK-----ELEKDPVWKKLVAVKQ 296
Db 227 EYPQ--YAQLSVERIIEANPDVVIYLIITH---GDPNSVKKAFEGEMMKNEAKNLDVAKQN 281
QY 297 RVDILDRDLWARSGLISSEEMAKELVELSKKDKK 332
Db 282 RVILPPDLFGSNPGTKVTEAM--DFMYKSIQDVKK 315

RESULT 34
US-09-071-035-40
; Sequence 40, Application US/09071035
; Publication No. US20020045737A1
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 40:
```

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 296 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-071-035-40

Query Match      15.9%; Score 268; DB 9; Length 296;
Best Local Similarity 30.3%; Pred. No. 2.8e-10;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

Qy 26 CGNN---SSNSSKSSKSGVEIKHEEGTTKVPKPKRVVVLVSFVDALVALDV--KPV 80
Db 1 CQTNKTADSAITETAKTEVTVDKTNGLTVPKPKRVVVDNGSLDMDALGVGDRV 60

Qy 81 GIADNKKNRILKPLRDKIGKYSVGTGRKQPNLEISKLKPDLLIADNNRHKGIYKDLNK 140
Db 61 GAPTGN-----IPAYLKYYQKVESAGGIKEPDLEKINQLKPDLLII--SGRQDDYQEQQLKA 114

Qy 141 IAPTIEL-----KSPDGDYNNENIDAFKTSKALGKEEKGKLEHDKKIEYKKEITMDK 196
Db 115 IAPTIYLAVDKPNWASTKQNIETLGTI---FDKEEVAKEKITGLEKEIADVKKQAEASA 171

Qy 197 NQKVLPAVAKSGLLAHPNSVVGQFLSQLGFKALSDD----VTKGLSKYLKGPYLOMN 252
Db 172 N-NALVVLVNEGQLSAYGKGRFGLIHDTFGKAA--DDKTEASTHGQS-----VS 219

Qy 253 TETLSQVNPERMFIM-TNKASSNEPSLKELEKDPVWKKLNAVQNRVDILDRDLWARSRG 311
Db 220 YEYVLEKNPGLFVVDRTKALGGDDSKDNVAANELIOKTDAGKDKVIMLPQDVWYLSGG 279

Qy 312 LISSEEMAKELVELSKK 328
Db 280 GLESMHL---MIEDVKK 293

RESULT 35
US-10-206-576-40
; Sequence 40, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
```

```
;
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 296 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-206-576-40

Query Match      15.9%; Score 268; DB 14; Length 296;
Best Local Similarity 30.3%; Pred. No. 2.8e-10;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

Qy 26 CGNN---SSNSSKSSKSGVEIKHEEGTTKVPKPKRVVVLVSFVDALVALDV--KPV 80
Db 1 CQTNKTADSAITETAKTEVTVDKTNGLTVPKPKRVVVDNGSLDMDALGVGDRV 60

Qy 81 GIADNKKNRILKPLRDKIGKYSVGTGRKQPNLEISKLKPDLLIADNNRHKGIYKDLNK 140
Db 61 GAPTGN-----IPAYLKYYQKVESAGGIKEPDLEKINQLKPDLLII--SGRQDDYQEQQLKA 114

Qy 141 IAPTIEL-----KSPDGDYNNENIDAFKTSKALGKEEKGKLEHDKKIEYKKEITMDK 196
Db 115 IAPTIYLAVDKPNWASTKQNIETLGTI---FDKEEVAKEKITGLEKEIADVKKQAEASA 171

Qy 197 NQKVLPAVAKSGLLAHPNSVVGQFLSQLGFKALSDD----VTKGLSKYLKGPYLOMN 252
Db 172 N-NALVVLVNEGQLSAYGKGRFGLIHDTFGKAA--DDKTEASTHGQS-----VS 219

Qy 253 TETLSQVNPERMFIM-TNKASSNEPSLKELEKDPVWKKLNAVQNRVDILDRDLWARSRG 311
Db 220 YEYVLEKNPGLFVVDRTKALGGDDSKDNVAANELIOKTDAGKDKVIMLPQDVWYLSGG 279

Qy 312 LISSEEMAKELVELSKK 328
Db 280 GLESMHL---MIEDVKK 293

RESULT 36
US-10-912-362-40
; Sequence 40, Application US/10912362
; Publication No. US20050043528A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 14200 Shady Grove Road
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 2000
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/912,362
; FILING DATE: 06-Aug-2004
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 10/206,576
; FILING DATE: 2002-07-29
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
```

```
;
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D2
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-912-362-40

Query Match 15.9%; Score 268; DB 17; Length 296;
Best Local Similarity 30.3%; Pred. No. 2.8e-10;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

QY 26 CGNN---SSNSKSSKSGVVEIKHEEGTTKVPKPKRVVVVLEYSFVDALVDV--KPV 80
Db 1 CQTKKTTADSAETTTAKTEVTITQNGQLTVPKPKRVVVFDNGLSLTMDALGVGDRVV 60

QY 81 GIADNKNRIIKPLRDKIKGYTSVGTQKPNLEIEISKLPDLIIADNNRHKGIIYKDLNK 140
Db 61 GAPTKN-----IPAYLKVKYVESAGSAGIKPEPLEKINQLKPDLLII-SGRQDQYQEQLKA 114

QY 141 IAPTIEL-----KSGFDGYNENIDAFKTIKALGKEEGKRLKEHDKKIEEYKKEITMDK 196
Db 115 IAPTIYLVADAKNPWASTKQNIETLGTI---PDKEEVAKEKITGLEKEIADYVKQAEASA 171

QY 197 NQKVLPAVAKSGLLAHPNSYVGQFLSQLGKFEALSDP-----VTKGLSKYLKGPYLOMN 252
Db 172 N-NALVVLVNEGQLSAYGKSRFGLIHDFTGFKAA--DDKIEASTHGQS-----VS 219

QY 253 TETLSQVNPFRMFI--TNKASSNEPSLKELEKDPVWKKNLNAVQNQRVDILDRDLWARSRG 311
Db 220 YEIVLEKNPGLFVVDRTKAIAGDDSKDNVAANELIKQTDAGKNDKVIMLQPDVVVYLSG 279

QY 312 LISSEEMAKELVELSKK 328
Db 280 GLESMHL---MIEDVKK 293

RESULT 37
US-09-738-626-4403
; Sequence 4403, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4403
; LENGTH: 306

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4403

Query Match 15.7%; Score 266; DB 9; Length 306;
Best Local Similarity 25.9%; Pred. No. 4e-10;
Matches 82; Conservative 60; Mismatches 135; Indels 40; Gaps 7;

QY 17 LFLVIATA---ACGNSSSSNSKSSKSGVVEIKHEEGTTKVPKPKRVVVVLEYSFVDAL 72
Db 11 LIVLAAALPLTACSSSESEASTSATR--EFTDAHGTTEVPENPQKRVVLEPLELDTA 68

QY 73 VALDVKPVGIADNKNRIIKPLRDKIKGYTSVGTQKPNLEIEISKLPDLIIADNNRHK 132
Db 69 IALGITPVGAANNVTGI--PAYLGVGDIGIEPVGTSEPNIEAIAALEPDLILGTDSRHA 126

QY 133 GIYKDLANKIAPTIELKSPFDGYNENIDAFKTIKALGKEEGKRLKEHDKKIEEYKKEI 192
Db 127 EYDRLESIAPTVFMTHVDPKQNV---VFIGALGKKQSEDLIQGFNDKCEIKSEH 183

QY 193 TMDKNQKVLPAVAKSGLLAHPNSYVGQFLSQLGKFEALSDVDVTGKLSKYLKGPYLOMN 252
Db 184 DVEGKTVNMIRPDEQTMSLYGPSTFAGSSLEACAG-----LTIPOQEWK 227

QY 253 TETLSQVNPFRMFI-----MTNKASSNEPSLKELEKDPVWKKNLNAVQNQRVDILDRDL 305
Db 228 DDLQADIAPENFMLATADYVFVTATDVTDENELPEVIRE-----NRQFPFSLTLVDTSY 281

QY 306 WARSGLISSEEMAKEL 322
Db 282 WVGVGPLGGSGKVLEDI 298

RESULT 38
US-10-627-476-572
; Sequence 572, Application US/10627476
; Publication No. US20040030116A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Mark
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schoder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CPCN
; CURRENT APPLICATION NUMBER: US/10/627,476
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 09/602,787
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 572
; LENGTH: 306
```



```

Qy 175 -----KGRLEEHKKIIEYKRTMTDNQKQVLPV-----AAKSGLLAHPSNSYVQ 221
Db 181 ELIKNTEARISKELEKH-----PEIKGKI-----KGKKVLFMTINAADTSKFWIYTSKDPDPRAN 233
Qy 222 FLSQLG--FKGALSDDVTGKLSKYLKGPYLQWNTETLSQVNPFRFIMTNKASSNEPSLK 279
Db 234 YLTDLGLVPFSPSLKEFESE--DSFAK-----EISAEENKINDADVII-----TYGDDKTLE 283
Qy 280 ELEKDPVWKKLNAVKNQVRDIL 301
Db 284 ALQKDPLLGKINAIKNGAVVI 305

```

Search completed: August 25, 2005, 09:40:02
Job time : 166 secs

This Page Blank (usptol)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:18:46 ; Search time 167 Seconds
(without alignments)
775.837 Million cell updates/sec

Title: US-10-724-972A-6352

Perfect score: 1690

Sequence: 1 GVESVRGLKILSVIGLFLVL.....EEMAKELVELSKDSKKDNK 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 16Dec04:*

- 1: Geneseqp1980a:*
- 2: Geneseqp1990a:*
- 3: Geneseqp2000a:*
- 4: Geneseqp2001a:*
- 5: Geneseqp2002a:*
- 6: Geneseqp2003a:*
- 7: Geneseqp2003b:*
- 8: Geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1690	100.0	335	8	ADs07057 Staphyloc
2	1668	98.7	331	6	ABU43214 Protein e
3	1242	73.5	337	4	Aau36661 Staphyloc
4	1233	73.0	337	6	ABR41865 Staphyloc
5	1232	72.9	327	6	ABU15904 Protein e
6	1228	72.7	337	6	ABR38894 Surface e
7	1214	71.8	348	6	ABM73426 Staphyloc
8	1181	69.9	309	4	Aau33963 Staphyloc
9	726	43.0	319	4	Aau33753 Staphyloc
10	726	43.0	319	4	Aau36529 Staphyloc
11	724.5	42.9	319	6	ABJ19052 Pathogen
12	724.5	42.9	319	6	ABU16209 Protein e
13	724.5	42.9	325	6	ABR38895 Surface e
14	724.5	42.9	325	6	ABR41867 Staphyloc
15	724.5	42.9	325	6	ABM72855 Staphyloc
16	644	38.1	175	6	ABU43367 Protein e
17	516.5	30.6	298	6	ABU38966 Protein e
18	476	28.2	303	6	ABM68354 Photorhab
19	460	27.2	300	6	ABU28130 Protein e
20	460	27.2	302	4	Aau34876 E. coli c
21	460	27.2	302	6	ABU28881 Protein e
22	454	26.9	297	6	ABU41496 Protein e
23	405	24.0	330	2	AAW89745 Staphyloc
24	405	24.0	330	2	AAV42331 Staphyloc
25	405	24.0	330	2	AAV31824 Staphyloc

26	405	24.0	330	6	ABJ19069 Pathogen
27	405	24.0	330	6	ABR38893 Surface e
28	405	24.0	330	6	ABR41864 Staphyloc
29	405	24.0	330	7	ADF43556 Staphyloc
30	391	23.1	324	6	ABU17675 Protein e
31	375	22.2	1132	4	ABG08851 Novel hum
32	350	20.7	932	4	ABG13148 Novel hum
33	343	20.3	287	6	ABM73050 Staphyloc
34	331	19.6	315	4	AGG90121 C glutami
35	327	19.3	321	6	ABU25343 Protein e
36	313	18.5	314	6	ABU18359 Protein e
37	307.5	18.2	352	7	ADC94889 E. faeciu
38	301	17.8	314	6	ABU17461 Protein e
39	300	17.8	264	6	ABU18269 Protein e
40	296.5	17.5	351	5	ABP40647 Staphyloc
41	296.5	17.5	351	8	ADS07886 Staphyloc
42	296	17.5	342	6	ABU40861 Protein e
43	296	17.5	351	7	ADF05577 Bacterial
44	292	17.3	347	4	ASG82087 S. epider
45	292	17.3	347	4	ASG81684 S. epider

ALIGNMENTS

RESULT 1

ADS07057
ID ADS07057 standard; protein; 335 AA.

XX ADS07057;

AC ADS07057;

DT 04-NOV-2004 (first entry)

DE Staphylococcus epidermis polypeptide seqid 6352.

XX antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW recombinant expression vector; infection; computer readable medium;
KW computer based system.

OS Staphylococcus epidermidis.

XX US2004147734-A1.

PN 29-JUL-2004.

PD 01-DEC-2003; 2003US-00724972.

PF 08-NOV-1997; 97US-0064964P.

PR 13-AUG-1998; 98US-00134001.

XX 29-NOV-1999; 99US-00450969.

PA (DOUC/) DOUCETTE-STAMM L.

XX (BUSH/) BUSH D.

PI Doucette-Stamm L, Bush D;

XX WPI: 2004-580138/56.

DR N-PSDB; ADS03285.

XX New isolated polypeptide and encoding nucleic acid derived from

PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or

PT treating an S. epidermidis bacterial infection.

XX Claim 17; SEQ ID NO 6352; 741pp; English.

PS The invention describes an isolated nucleic acid comprising a nucleotide

XX sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:

CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any

CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as

CC given in the specification. Also described are: a recombinant expression

CC vector; a cell comprising a recombinant expression vector of (1);

CC producing an S. epidermidis polypeptide; an isolated nucleic acid

CC comprising a nucleotide sequence of at least 8 nucleotides in length; a

CC vaccine composition for prevention or treatment of an *S. epidermidis*
 CC infection, comprising a nucleic acid cited above and a carrier; treating
 CC a subject for *S. epidermidis* infection; a recombinant or substantially
 CC pure preparation of an *S. epidermidis* polypeptide or its fragment; a
 CC vaccine composition for prevention or treatment of an *S. epidermidis*
 CC infection; detecting the presence of a *Staphylococcus* nucleic acid in a
 CC sample; a computer readable medium having recorded in it the nucleotide
 CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
 CC system for identifying fragments of the *Staphylococcus* genome of
 CC commercial importance; a computer based system for identifying fragments
 CC of the *Staphylococcus* plasmids of commercial importance; identifying
 CC commercially important nucleic acid fragments of the *Staphylococcus*
 CC genome and/or plasmids; and identifying an expression modulating fragment
 CC of the *Staphylococcus* genome and/or plasmids. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of an *Staphylococcus* epidermidis bacterial
 CC infection. This is the amino acid sequence of a *S. epidermis* protein of
 CC the invention.

XX Sequence 335 AA;

Query Match 100.0%; Score 1690; DB 8; Length 335;
 Best Local Similarity 100.0%; Pred. No. 1.5e-129;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVESVRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 60
 DB 1 GVESVRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 60

QY 61 VVLYSFVDALVALDVPKVGADDNKNRIIKPLRDKIGKTVSGTRKQPNLEISKLK 120
 DB 61 VVLYSFVDALVALDVPKVGADDNKNRIIKPLRDKIGKTVSGTRKQPNLEISKLK 120

QY 121 PDLIADNNRHKGIIYKDLNKTAPIETLKSFDGDYNNENIDAFKTIISKALGKEEGKRLKEE 180
 DB 121 PDLIADNNRHKGIIYKDLNKTAPIETLKSFDGDYNNENIDAFKTIISKALGKEEGKRLKEE 180

QY 181 HDKKEEYKEITMDKNQKVLPAVAAGSGLLAHPNSVYVGQFLSQLGPKAALSDVTYKGL 240
 DB 181 HDKKEEYKEITMDKNQKVLPAVAAGSGLLAHPNSVYVGQFLSQLGPKAALSDVTYKGL 240

QY 241 SKYLKGPYLQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKKLNVAQNQRVDI 300
 DB 241 SKYLKGPYLQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKKLNVAQNQRVDI 300

QY 301 LDRDLWARSRLGISSEEMAKELVELSKDKDNK 335
 DB 301 LDRDLWARSRLGISSEEMAKELVELSKDKDNK 335

RESULT 2

ABU43214
 ID ABU43214 standard; protein; 331 AA.

AC ABU43214;

DT 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #28741.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX *Staphylococcus epidermidis*.

OS WO20027183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-0009107.

PF 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA47084.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 71138; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid;
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 331 AA;

Query Match 98.7%; Score 1668; DB 6; Length 331;
 Best Local Similarity 99.7%; Pred. No. 9.2e-128;
 Matches 330; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 VRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKRVVVL 64
 DB 1 MRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKRVVVL 60

QY 65 EYSFVDALVALDVPKVGADDNKNRIIKPLRDKIGKTVSGTRKQPNLEISKLKPDLI 124

DB 61 EYSFVDALVALDVPKVGADDNKNRIIKPLRDKIGKTVSGTRKQPNLEISKLKPDLI 120

QY 125 IADNNRHKGIIYKDLNKTAPIETLKSFDGDYNNENIDAFKTIISKALGKEEGKRLKEEDKK 184

DB 121 IADNNRHKGIIYKDLNKTAPIETLKSFDGDYNNENIDAFKTIISKALGKEEGKRLKEEDKK 180

QY 185 IEEYKKEITMDKNQKVLPAVAAGSGLLAHPNSVYVGQFLSQLGPKAALSDVTYKGLSKYL 244

DB 181 IEEYKKEITMDKNQKVLPAVAAGSGLLAHPNSVYVGQFLSQLGPKAALSDVTYKGLSKYL 240

QY 245 KGPYLQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKKLNVAQNQRVDILDRD 304

KW SEIP; surface exposed immunogenic polypeptide; iron-binding; receptor;
 KW antibacterial; fungicide; protozoacide; infection; disinfection.
 XX Unidentified.
 OS WO2003006672-A2.
 PN 23-JAN-2003.
 XX
 XX
 XX 10-JUL-2002; 2002WO-US021663.
 XX
 XX 10-JUL-2001; 2001US-0304390P.
 XX
 XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
 XX
 XX Scott DL;
 XX
 XX WPI; 2003-221756/21.
 XX
 XX Isolating microbial target surface exposed immunogenic polypeptides, by
 PT immunizing host with microbial membrane-associated polypeptides to
 PT produce antibodies and using antibodies to isolate target polypeptides.
 XX
 XX Claim 15; Page 40-41; 46pp; English.
 XX
 XX The invention relates to the isolation of target surface exposed
 CC immunogenic polypeptides (SEIP) from microorganisms. The SEIP's of the
 CC invention are receptors of iron-binding molecules. The activity of
 CC polypeptides of the invention may be described as antibacterial,
 CC fungicide and protozoacide. Pharmaceutical compositions based on the
 CC polypeptides of the invention are useful for treating infections and
 CC diseases caused by microbes including bacteria, fungi and protozoa in a
 CC patient. They are also useful for topical disinfection in an animal or
 CC mammal e.g. human. The current sequence represents an SEIP of the
 CC invention
 XX
 XX Sequence 337 AA;
 SQ
 Query Match 72.7%; Score 1228; DB 6; Length 337;
 Best Local Similarity 71.3%; Pred. No. 8.4e-92;
 Matches 234; Conservative 43; Mismatches 51; Indels 0; Gaps 0;
 QY 4 SVRGLKILSVIGLLFVLIATAACGNSSNSSKSSKDGVEIKHEGTTKVPKPKRVVV 63
 DB 10 TWRGLKTFILGLIIVALLVAACGNTDSSKSSSTKDTISVKDENGTVKPKDAKRIIV 69
 QY 64 LEYSFVDALVALDVKPVGIIADNKKRIIIPKLRDKIGKTVTSVGTQKPNLEISKLPDL 123
 DB 70 LEYSFADALVALDVKPVGIIADNKKRIIIPKVRKIGDVTSGTRKQPNLEISKLPDL 129
 QY 124 IADNRRHGIYKDLNKIAPTIELKSPDGYNENIDAPKTIISKALGKEEGKLEEHDK 183
 DB 130 IADSSRRHGINKELNKIAPTIELKSPDGYNENIDAPKTIISKALGKEEGKLEEHDK 189
 QY 184 KIEEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVVGQFLSQGLFKEALSDVTKGLSKY 243
 DB 190 LINKYKDEIKFDRNQKVLPAVAKAGLLAHPNTSYVVGQFLNELGFKNALSDVTKGLSKY 249
 QY 244 LKGPYLQMNVTETLSQVNPFRMTNWKASNEPSLKEKDPVWKKNVAVQNRVDILDR 303
 DB 250 LKGPYLQDLTEHLADLNPERMIIMTDHAKKDSAEFFKQLQEDATWKKLVAVKNNRVDIVDR 309
 QY 304 DLWARSRLISSEEMAKELVELSKKOSK 331
 DB 310 DVWARSRLISSEEMAKELVELSKKQK 337
 RESULT 7
 ABM73426
 ID ABM73426 standard; protein; 348 AA.
 XX
 XX ABM73426;
 XX

DT 20-NOV-2003 (first entry)
 XX Staphylococcus aureus protein #2666.
 DE
 XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.
 XX
 XX Staphylococcus aureus.
 OS
 XX WO200294868-A2.
 PN
 XX 28-NOV-2002.
 PD
 XX 27-MAR-2002; 2002WO-IB002637.
 PF
 XX 27-MAR-2001; 2001GB-00007661.
 PR
 XX (CHIR-) CHIRON SPA.
 XX
 XX Masignani V, Mora M, Scarselli M;
 PI
 XX WPI; 2003-120786/11.
 DR
 XX N-PSDB; ACF74986.
 XX
 XX New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.
 XX
 XX Claim 1; SEQ ID NO 5332; 49pp; English.
 PS
 XX The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus proteins of the invention
 XX
 XX Sequence 348 AA;
 SQ
 Query Match 71.8%; Score 1214; DB 6; Length 348;
 Best Local Similarity 71.8%; Pred. No. 1.2e-90;
 Matches 232; Conservative 41; Mismatches 50; Indels 0; Gaps 0;
 QY 9 KILSVIGLLFVLIATAACGNSSNSSKSSKDGVEIKHEGTTKVPKPKRVVLEYSF 68
 DB 26 KPSILGLIIVALLVAACGNTDSSKSSSTKDTISVKDENGTVKPKDAKRIIVLEYSF 85
 QY 69 VDVALVDVKPVGIIADNKKRIIIPKLRDKIGKTVTSVGTQKPNLEISKLPDLIADN 128
 DB 86 ADALALVDVKPVGIIADNKKRIIIPKVRKIGDVTSGTRKQPNLEISKLPDLIADN 145
 QY 129 NRHGIYKDLNKIAPTIELKSPDGYNENIDAPKTIISKALGKEEGKLEEHDKKIEEY 188
 DB 146 SRHGINKELNKIAPTIELKSPDGYNENIDAPKTIISKALGKEEGKLEEHDKKIEEY 205
 QY 189 KKEITMDKNQKVLPAVAKSGLLAHPNSYVVGQFLSQGLFKEALSDVTKGLSKYKGPY 248
 DB 206 KDEIKFDRNQKVLPAVAKAGLLAHPNTSYVVGQFLNELGFKNALSDVTKGLSKYKGPY 265
 QY 249 LQMNVTETLSQVNPFRMTNWKASNEPSLKEKDPVWKKNVAVQNRVDILDRDLW 308
 DB 266 LQDLTEHLADLNPERMIIMTDHAKKDSAEFFKQLQEDATWKKLVAVKNNRVDIVDRV 325
 QY 309 SRGLISSEEMAKELVELSKKOSK 331
 DB 326 SRGLISSEEMAKELVELSKKQK 348
 RESULT 8


```

AAU33963
ID AAU33963 standard; protein; 309 AA.
XX
AC AAU33963;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #239.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
XX
PR 23-MAY-2000; 2000US-0206848P.
XX
PR 26-MAY-2000; 2000US-0207727P.
XX
PR 23-OCT-2000; 2000US-0242578P.
XX
PR 27-NOV-2000; 2000US-0253625P.
XX
PR 22-DEC-2000; 2000US-0257931P.
XX
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
WPI; 2001-611495/70.
DR N-PSDB; AAS51822.
XX
PT New polynucleotides for the identification and development of
XX
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 5459; 51lpp; English.
XX
SQ The invention relates to antisense inhibitors of genes essential to
XX
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX
XX their use in the discovery of novel antibiotics, the essential genes,
XX
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX
XX useful for the identification of potential new targets for antibiotic
XX
XX development. The antisense nucleic acids can also be used to identify
XX
XX proteins used in proliferation, to express these proteins, and to obtain
XX
XX antibodies capable of binding to the expressed proteins. The proteins can
XX
XX be used to screen compounds in rational drug discovery programmes. The
XX
XX antisense nucleic acid sequence is also useful to screen for homologous
XX
XX nucleic acids which are required for cell proliferation in a wide variety
XX
XX of organisms. The present sequence represents an essential prokaryotic
XX
XX cellular proliferation protein. Note: The sequence data for this patent
XX
XX did not form part of the printed specification, but was obtained in
XX
XX electronic format directly from WIPO at
XX
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 309 AA;
XX
Query Match 69.9%; Score 1181; DB 4; Length 309;
Best Local Similarity 72.8%; Pred. No. 5.2e-88;
Matches 225; Conservative 38; Mismatches 46; Indels 0; Gaps 0;
XX
QY 16 LLFVLIAACGNNSSSSKSSKDGVEIKHECTTKVPHKRVVVLVYSFVDAVAL 75
DB 1 LIFVLIAACGNTDNSSKSSSTKDTISVDKENGTVKVPDAAKRVVLVYSFADALAA 60
QY 76 DVKPGVIADNKKRIIPLRDKIKGYTSVGRKOPNLEEISKLKPDLIADNNRHKGIV 135
DB 61 DVKPGVIADGKKRIIPVREKIGNYTSVGRKOPNLEEISKLKPDLIADSSRHKGIN 120

```

```

QY 136 KDLNKIAPTIELKSFQDGYNNENIDAFKTIISKALGKBEGKKRLREHDKKIIEYKKEITMD 195
DB 121 KELNKIAPTISLSKSFQDGYKQNIINSFKITIAALNKEKEGKRLAEHDKLIKKYKDEIKFD 180
QY 196 KNQKVLPAVAAKSGLLAHPNSYVYGQFLSQGLQFKKALSDDDVTGKLSKYLKGPYLOMNTET 255
DB 181 RNQKVLPAVAAKAGLLAHPNYSYVYGQFLNELGFKNALSDDDVTGKLSKYLKGPYLOMNTET 240
QY 256 LSQVNPFRMFTMTNKASNEPSLKELEKDPWKKLNAVKNORVDILDRDLWARSGLISS 315
DB 241 LADLNPERMTIIMTDNAKKDSAEFKKLOEDPTWKKLNAVKNRVDIVDRDWARSGRLISS 300
QY 316 EEMAKELVE 324
DB 301 EEMAKELVE 309
XX
RESULT 9
AAU33753
ID AAU33753 standard; protein; 319 AA.
XX
AC AAU33753;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #29.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
XX
PR 23-MAY-2000; 2000US-0206848P.
XX
PR 26-MAY-2000; 2000US-0207727P.
XX
PR 23-OCT-2000; 2000US-0242578P.
XX
PR 27-NOV-2000; 2000US-0253625P.
XX
PR 22-DEC-2000; 2000US-0257931P.
XX
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
WPI; 2001-611495/70.
DR N-PSDB; AAS51612.
XX
PT New polynucleotides for the identification and development of
XX
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 5249; 51lpp; English.
XX
SQ The invention relates to antisense inhibitors of genes essential to
XX
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX
XX their use in the discovery of novel antibiotics, the essential genes,
XX
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX
XX useful for the identification of potential new targets for antibiotic
XX
XX development. The antisense nucleic acids can also be used to identify
XX
XX proteins used in proliferation, to express these proteins, and to obtain
XX
XX antibodies capable of binding to the expressed proteins. The proteins can
XX
XX be used to screen compounds in rational drug discovery programmes. The
XX
XX antisense nucleic acid sequence is also useful to screen for homologous
XX
XX nucleic acids which are required for cell proliferation in a wide variety

```


PA (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
DR WPI; 2003-029926/02.
DR N-PSDB; ACA47237.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX
PS Claim 25; SEQ ID NO 71291; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 175 AA;
Query Match 38.1%; Score 644; DB 6; Length 175;
Best Local Similarity 70.9%; Pred. No. 1.8e-44;
Matches 124; Conservative 23; Mismatches 24; Indels 4; Gaps 1;
QY 5 VRGLKILSVIGLFLVLIATAACGNSNSNS----SKESKDGVEIKHEGTTKVPKPKR 60
DB 1 MKGFRPAGIVALLFALVLTACGNSNGSGDSKSSKSIKIKHELGTTKVPKDKR 60
QY 61 VVVEFSFDALVALDVKPGVADNKKRIIPLRDKTKGTVSVTRQPNLEISKUK 120
DB 61 VVVEFSFDALVALDVKPGVADNKKRIIPLRDKTKGTVSVTRQPNLEISKUK 120
QY 121 PLLIADNNRHGIYKDLNKAFTIELKSFDDGYNENIDAFKTIKALCKEKGK 175
DB 121 PLLIADNNRHGIYKDLNKAFTIELKSFDDGYNENIDAFKTIKALCKEKGK 175

RESULT 17
ABU38966
ID ABU38966 standard; protein; 298 AA.
XX
AC ABU38966;
XX
DT 19-JUN-2003 (first entry)
XX

DE Protein encoded by Prokaryotic essential gene #2493.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Pasteurella multocida.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI WPI; 2003-029926/02.
DR N-PSDB; ACA42836.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 66890; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 298 AA;
Query Match 30.6%; Score 516.5; DB 6; Length 298;
Best Local Similarity 38.1%; Pred. No. 9.5e-34;
Matches 111; Conservative 65; Mismatches 98; Indels 17; Gaps 4;
QY 43 VEIKHEGTTKVPKPKRVVLEYSFDALVALDVKPGVADNKKRIIPLRDKIGKY 102
DB 23 VTVKQKGFTLDSVPKRVVALEYSVDALAIGVSPGVGADNDKTRILQKVRDKVQPW 82

the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid; (2) a host cell containing the vector; (3) an isolated nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 300 AA;

Query Match 27.2%; Score 460; DB 6; Length 300;
Best Local Similarity 35.8%; Pred. No. 4e-29;
Matches 101; Conservative 60; Mismatches 105; Indels 16; Gaps 4;
QY 45 IKHEGTTKPKPRVVLVLEYSFVDALVALDVKPVGIADNNKKNRIKPLRDKIGKYS 104
DB 24 VQDEHGTFTLEKTPQRIIVVLELSPADALAAVDVPIGIIADNDKAKRILPEVRAHLKPMQS 83
QY 105 VTRKQPNLEISKLKPDLLIADNNRHGIYKLNKIAPTIELKSGFDGYNENIDAFKTI 164
DB 84 VGTQAQPSLEATAALKPDLLIADSSRHAGIYVIALQOIAPVLLKSRNETYAEINLSAAII 143
QY 165 SKALGKEEGKKRLEHDKKIEYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVQGLS 224
DB 144 GEVVGKKREMQARLEQHKERMAQWSSQ--LPTGTRVAFGTSREQLNLHTQETWTGSLA 201
QY 225 QLGFKREALSDVTKGLSKYLKGPYL-QMNTETLSQVNPMPFIMTNKASSNEPFLKLEBK 283
DB 202 SLGL-----NVPAMAGASMPISGLEQLLAVNPAILVAHYR---EESIVKRWQ 248
QY 284 DPVWKKLNKQVQVVDILDRDLWARSRLISSEEMAKELVEL 325
DB 249 DPLWQMLTAAQKQVQVSDNSNWARMRGIFAAERIAADTVKI 290

RESULT 20
AAU34876
ID AAU34876 standard; protein; 302 AA.
XX
AC AAU34876;
XX
DT 14-FEB-2002 (first entry)
XX
DE E. coli cellular proliferation protein #457.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Escherichia coli.
XX

PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS52735.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 10469; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 302 AA;
Query Match 27.2%; Score 460; DB 4; Length 302;
Best Local Similarity 35.5%; Pred. No. 4e-29;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;
QY 45 IKHEGTTKPKPRVVLVLEYSFVDALVALDVKPVGIADNNKKNRIKPLRDKIGKYS 104
DB 26 VQDEHGTFTLEKTPQRIIVVLELSPADALAAVDVPIGIIADNDKAKRILPEVRAHLKPMQS 85
QY 105 VTRKQPNLEISKLKPDLLIADNNRHGIYKLNKIAPTIELKSGFDGYNENIDAFKTI 164
DB 86 VGTQAQPSLEATAALKPDLLIADSSRHAGIYVIALQOIAPVLLKSRNETYAEINLSAAII 145
QY 165 SKALGKEEGKKRLEHDKKIEYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVQGLS 224
DB 146 GEVVGKKREMQARLEQHKERMAQWSSQ--LPTGTRVAFGTSREQLNLHTQETWTGSLA 203
QY 225 QLGFKREALSDVTKGLSKYLKGPYL-QMNTETLSQVNPMPFIMTNKASSNEPFLKLEBK 283
DB 204 SLGL-----NVPAMAGASMPISGLEQLLAVNPAILVAHYR---EESIVKRWQ 250
QY 284 DPVWKKLNKQVQVVDILDRDLWARSRLISSEEMAKELVEL 325
DB 251 DPLWQMLTAAQKQVQVSDNSNWARMRGIFAAERIAADTVKI 292

```
RESULT 21
ABU28881.
ID ABU28881 standard; protein; 302 AA.
XX AC ABU28881;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #14408.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Escherichia coli.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX FA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA32751.
XX DE New antisense nucleic acids, useful for identifying proteins or screening
XX DE for homologous nucleic acids required for cellular proliferation to
XX DE isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 56805; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: the sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX CC Sequence 302 AA;
Query Match 27.2%; Score 460; DB 6; Length 302;
Best Local Similarity 35.5%; Pred. No. 4e-29;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;
QY 45 IKHEEGTTKVPKPKRVVLEYSFVDALVALDVKVPVGADNNKKNRIKPLRDKIGKYTS 104
DB 26 VODEHGTFLEKTPQRIIVLELSPADALAAVDVIPIGIADDNDKAKILPEVRAHLKPQWS 85
QY 105 VGTTRKQNLBEISKLPDIIADNNRHKGIIKDLINKIAPTIELKSFDDYNNENIDAFKTI 164
DB 86 VGTRAQPSLEAIAALKPDIIADSSRHAGVYIALQQIAPVLLKSRNETYAENLQSAAIL 145
QY 165 SKALGKEEGKRLLEHDKKIEEYKKEITMDKNOKVLPAAVAKSGLLAHPNSVYQFLS 224
DB 146 GEMVKKREMQARLEQHKERMAQWASQ--LPKGRVAFGTSREQQFNLHTQETWTGSLA 203
QY 225 QLGFEALSDDDVTGKSLKYLKPYL-QMNTETLSQVNPFRPIMTNKASSNEPSLKELEK 283
DB 204 SLGL-----NVPAAMAGASMPSTIGLEQLLAVNPALLVAHYR---EESIVKRWQQ 250
QY 284 DPVWKKNLNAVKNQVDIILDRDLWARSGLISSSEMAKELVEL 325
DB 251 DPLWQMLTAAQKQOVASVDSNTWARMRGIFAAERIAADTVKI 292
RESULT 22
ABU41496
ID ABU41496 standard; protein; 297 AA.
XX AC ABU41496;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #27023.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Pseudomonas syringae.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX FA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA45366.
XX DE New antisense nucleic acids, useful for identifying proteins or screening
XX DE for homologous nucleic acids required for cellular proliferation to
XX DE isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 69420; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
```

antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 297 AA;

Query Match 26.9%; Score 454; DB 6; Length 297;

Best Local Similarity 34.3%; Pred. No. 1.2e-28;

Matches 109; Conservative 60; Mismatches 109; Indels 40; Gaps 6;

QY 17 LFLVLTATACGNNSSNSSKESKDGVEIKHEEGTTKV--PKHPRVVVLEYSFVDALVA 74
 DB 5 LTLTASAA-----QAAPIDIDGQHKVLPDTPPKRVVVLEFSFIDGLAS 49
 QY 75 LDVQVPGIADNKNRIIKPLRDKIGKVTSGTRKQPNLEETSKLPDIIADNNRHKG1 134
 DB 50 VGVTPGAADDGASRVLPKRVKAVGWSVGLRQPNIEVTLARKLPDIIADLGRHQAL 109
 QY 135 YKDLNKIAPTIELKSGFDGYNENIDAFKTI SKALGKEEGKRLBEHKKI EYKBITM 194
 DB 110 YNDLASLAPTLPSRGEDYQGLSKSAGLIGMALGKGPQEMQARIAENRQHLKTVAEQIPA 169
 QY 195 DGNQKVLPAVAKSGLLAHPNSYVQFLSQFLSGFKEALSDDVTYKGLSKYLKPYLQMTTE 254
 DB 170 DSN--VLFVGAREDSFVGHPSYAGSVLQATG-----LQVPEVRNNA 211
 QY 255 TLSQVNPFRMFTWTK-----ASSNEPSLKEL-EKDPVWKKNLAVKNQVRVDIILDRILWARS 309
 DB 212 PTEFVSLEQLALDPNWLIVGHYRRPSIVDTWSKQPLWQLVGLGNKQVAEVDGDSWARN 271
 QY 310 RGLISEEMAKELVELSK 327
 DB 272 RGMASEQIADALAVLK 289

RESULT 23

AAW89745

ID AAW89745 standard; protein; 330 AA.

AAW89745;

16-MAR-1999 (first entry)

Staphylococcus aureus protein SEQ ID #5193.

Computer readable medium; vaccine; *S. aureus* infection; immunodetection; cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy; skin infection; surgical wound infection; scalded skin syndrome; toxic shock syndrome.

Staphylococcus aureus.

PN EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -

XX stored on computer readable medium and used in the production of anti-

XX *S. aureus* vaccines.

XX Claim 23; Page 3189-3190; 3271pp; English.

XX This sequence represents a *Staphylococcus aureus* protein sequence of the

XX invention. The DNA sequences encoding the *S. aureus* proteins are recorded

XX on a computer readable medium, preferably selected from a floppy or hard

XX disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.

XX Homology searches using the *S. aureus* DNA sequences allows putative

XX functions to be assigned so that protein-encoding or regulatory regions

XX of commercial, therapeutic or industrial importance can be obtained.

XX Specifically, sequences which are likely to encode antigens have been

XX identified and these polypeptides can be used in a vaccine composition

XX against *S. aureus* infection. The polypeptides can also be used in a kit

XX for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated

XX in numerous human diseases, including cellulitis, eyelid infections, food

XX poisoning, osteomyelitis, skin and surgical wound infections, scalded

XX skin syndrome, toxic shock syndrome, etc. Organisms transformed with the

XX DNA sequences can be used for recombinant production of the polypeptides.

XX The new DNA sequences (and their fragments) are useful as primers or

XX probes for isolating homologues of any of the 5191 *S. aureus* DNA sequences

XX contained on the computer readable medium

XX Sequence 330 AA;

XX Query Match 24.0%; Score 405; DB 2; Length 330;

XX Best Local Similarity 32.8%; Pred. No. 1.4e-24;

XX Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY 8 LKILSVIGLLFVLITATAACGNNSSNSSKESKDGVEIKHEEGTTKVPKHPKRVVVLEYS 67

DB 5 IKML-VVTLAFLV-LAGCGSNKQSSDNKDKETTSIKHAMGTTTEIKGPKRVVTVLYQG 62

QY 68 FVDALVALDVKPVGIADNKNRIIKPLRDKIGKVTSGTRKQPNLEETSKLPDIIAD 127

DB 63 ATDVAVSLGVKPVGAVESWTQPKFEYIKNDLKDTKIVGQEPAPNLEETSKLPDIIAS 122

QY 128 NNRHKGIIYKDLNKIAPTIELKSGFDGYNENIDAFKTI SKALGKEEGKRLBEHKKIEE 187

DB 123 KVRNEKVVYDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEAEADLLKDYDDKVA 178

QY 188 YKKEITMDKNQKVLPAVAKSGLLAHPNS-----SYVGQFLSQFLGKEALSDDVTYKGL 240

DB 179 FQK----DAKAKYKDAWPLKASVNVNFRADHTRIYAGGAGETINDLGFKR-----NKDL 228

QY 241 SKYLAG--PYLQMTN-ETLSOVNPERMFTMTKASSNEPSL-----KELEKDPVWKKNLA 292

DB 229 QKQVNDGKDIIQLTSKESIPLMNADHIFVVKSDPNKADALVKKYTESSEWTSKWKQLDA 288

QY 293 VKNQVQVDDLDRIITWNLAGGYSKLIDDLVE--KLNIEKQSK 330

DB 289 VKNQVQVDDLDRIITWNLAGGYSKLIDDLVE--KLNIEKQSK 330

XX RESULT 24

XX AAY42331

ID AAY42331 standard; protein; 330 AA.

XX SEIP; surface exposed immunogenic polypeptide; iron-binding; receptor;
 KW antibacterial; fungicide; protozoacide; infection; disinfection.
 XX Unidentified.
 OS WO2003006672-A2.
 XX 23-JAN-2003.
 XX 10-JUL-2002; 2002WO-US021663.
 XX 10-JUL-2001; 2001US-0304390P.
 XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
 XX Scott DL;
 XX WPI; 2003-221756/21.
 XX Isolating microbial target surface exposed immunogenic polypeptides, by
 PT immunizing host with microbial membrane-associated polypeptides to
 PT produce antibodies and using antibodies to isolate target polypeptides.
 XX Claim 15; Page 40; 46pp; English.
 XX The invention relates to the isolation of target surface exposed
 CC immunogenic polypeptides (SEIP) from microorganisms. The SEIP's of the
 CC invention are receptors of iron-binding molecules. The activity of
 CC polypeptides of the invention may be described as antibacterial,
 CC fungicide and protozoacide. Pharmaceutical compositions based on the
 CC polypeptides of the invention are useful for treating infections and
 CC diseases caused by microbes including bacteria, fungi and protozoa in a
 CC patient. They are also useful for topical disinfection in an animal or
 CC mammal e.g. human. The current sequence represents an SEIP of the
 CC invention
 XX Sequence 330 AA;
 SQ
 Query Match 24.0%; Score 405; DB 6; Length 330;
 Best Local Similarity 32.8%; Pred. No. 1.4e-24;
 Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;
 QY 8 LKILSVIGLLPVLIATAACGNSSSSSKSSKDGVEIKHEEGTKVPKPKRVVVLEYS 67
 DB 5 IKML-VVTLAFLV-LAGCGSNKQSSDNKDKEKTTSIKHAMGTTEIKGPKRVVTLVYQ 62
 QY 68 FVDALVALDVPVGIADNKKRIIKPLRDKIGKYSVGTGRKQPNLEISKLKPDLLIAD 127
 DB 63 ATDVAVSLGVFPVGAVESWTQKPFYIKNDLKTKIVGQEPAPNLEISKLKPDLLIVAS 122
 QY 128 NNRHKGIIYKDLNKTAPTIELKSPFDGYNENIDAFKTSIKALGKEEGKRLKEEHDKIEE 187
 DB 123 KVRNEKYVDQLSKIAPTIV---STDVTFKFK-DTTKMGKALGKEAEADLLKKYDDKVA 178
 QY 188 YKKEITMDKNQKVLPAVAASKGLLAHPN-----SYVGQFSLQGFKEALSDDVTGKL 240
 DB 179 FQK----DAKAKYKDAWPLKASVNVFRADHTRIVAGGYAGEILNDLGFKR-----NKDL 228
 QY 241 SKYLGK--PYLQMT-ETLSQVNPFRFIMTNKASSNEPSL-----KELEKDPVWKKLNA 292
 DB 229 QKQVNDGKDIQLTSKESIPLMNADHIFVVKSDPNKADALVKKTESWTSSKEWKNLDA 288
 QY 293 VKNQRV-DILDRDLWARGSLISSEEMAKELVELSKDKSKDNK 335
 DB 289 VKNNQVSDLDLDEITWNLAGGYKSLKLDLDDLYE--KLNIEKQSK 330
 RESULT 28
 ABR41864
 ID ABR41864 standard; protein; 330 AA.
 XX
 AC ABR41864;

XX 11-AUG-2003 (first entry)
 XX Staphylococcal surface-exposed immunogenic polypeptide D2 SA02.
 DE Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
 KW antibacterial; vaccine.
 KW Staphylococcus aureus.
 OS WO2003020875-A2.
 XX 13-MAR-2003.
 XX 17-JUN-2002; 2002WO-US019224.
 XX 17-JUN-2001; 2001US-0298975P.
 XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
 XX Scott DL;
 XX WPI; 2003-300870/29.
 XX N-PSDB; ACC48531.
 XX Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
 PT containing receptors for siderophores or iron-binding ligands, useful for
 PT producing antibodies effective against Staphylococci infection.
 XX Claim 3; Page 51; 62pp; English.
 XX The present sequence is that of a novel Staphylococcus aureus surface-
 CC exposed immunogenic polypeptide (SEIP). To isolate SEIP genes, an
 CC expression library of S. aureus genomic DNA was screened with anti-SEIP
 CC antisera. DNA was isolated from clones expressing SEIPs, and amino acid
 CC sequences were deduced and analysed for conserved and/or functional
 CC domains. The present SEIP has sequence homology with the siderophore
 CC family of periplasmic binding proteins. The invention provides methods
 CC for the identification, production and recovery of SEIPs. The SEIPs can
 CC be used individually, or in combination, to produce anti-staphylococcal
 CC antibodies useful in passive or active immunisation strategies to prevent
 CC or contain staphylococcal infection. They can also be used to develop
 CC diagnostic assays
 XX Sequence 330 AA;
 SQ
 Query Match 24.0%; Score 405; DB 6; Length 330;
 Best Local Similarity 32.8%; Pred. No. 1.4e-24;
 Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;
 QY 8 LKILSVIGLLPVLIATAACGNSSSSSKSSKDGVEIKHEEGTKVPKPKRVVVLEYS 67
 DB 5 IKML-VVTLAFLV-LAGCGSNKQSSDNKDKEKTTSIKHAMGTTEIKGPKRVVTLVYQ 62
 QY 68 FVDALVALDVPVGIADNKKRIIKPLRDKIGKYSVGTGRKQPNLEISKLKPDLLIAD 127
 DB 63 ATDVAVSLGVFPVGAVESWTQKPFYIKNDLKTKIVGQEPAPNLEISKLKPDLLIVAS 122
 QY 128 NNRHKGIIYKDLNKTAPTIELKSPFDGYNENIDAFKTSIKALGKEEGKRLKEEHDKIEE 187
 DB 123 KVRNEKYVDQLSKIAPTIV---STDVTFKFK-DTTKMGKALGKEAEADLLKKYDDKVA 178
 QY 188 YKKEITMDKNQKVLPAVAASKGLLAHPN-----SYVGQFSLQGFKEALSDDVTGKL 240
 DB 179 FQK----DAKAKYKDAWPLKASVNVFRADHTRIVAGGYAGEILNDLGFKR-----NKDL 228
 QY 241 SKYLGK--PYLQMT-ETLSQVNPFRFIMTNKASSNEPSL-----KELEKDPVWKKLNA 292
 DB 229 QKQVNDGKDIQLTSKESIPLMNADHIFVVKSDPNKADALVKKTESWTSSKEWKNLDA 288
 QY 293 VKNQRV-DILDRDLWARGSLISSEEMAKELVELSKDKSKDNK 335
 DB 289 VKNNQVSDLDLDEITWNLAGGYKSLKLDLDDLYE--KLNIEKQSK 330


```
RESULT 29
ADF43556
ID ADF43556 standard; protein; 330 AA.
XX
AC ADF43556;
XX
DT 12-FEB-2004 (first entry)
XX
DE Staphylococcus aureus cbrA.
XX
KW antibacterial; immunostimulant; vaccine; vaccine;
KW Staphylococcus aureus infection; infection prevention;
KW infection attenuation; cbrA.
XX
OS Staphylococcus aureus.
XX
PN US2003153733-A1.
XX
PD 14-AUG-2003.
XX
PF 24-OCT-2002; 2002US-00278946.
XX
PR 20-MAR-1998; 98US-0078862P.
PR 01-APR-1998; 98US-0080296P.
PR 07-MAY-1998; 98US-0084674P.
PR 18-MAR-1999; 99WO-US006199.
PR 15-JAN-2002; 2002US-00830217.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Simpson AJG, Choi GH;
XX
DR WPI; 2003-897737/82.
DR N-PSDB; ADF43555.
XX
PT New Staphylococcus aureus gene, useful for preparing a vaccine against
PT infection caused by Staphylococcus aureus.
XX
PS Claim 9; SEQ ID NO 16; 61pp; English.
XX
CC The invention describes a new isolated nucleic acid molecule comprising:
CC a sequence that is 95% identical to a fully defined sequence having 586-
CC 2226 bp, or its complement; or a sequence encoding a polypeptide having a
CC fully defined sequence comprising 136-691 amino acids, or its complement.
CC The gene is Staphylococcus aureus gene. The nucleic acid is useful for
CC preparing a vaccine against infection caused by Staphylococcus aureus.
CC The methods are useful for preventing or attenuating an infection caused
CC by a Staphylococcus, detecting Staphylococcus nucleic acids in a
CC biological sample obtained from an animal, and detecting Staphylococcus
CC antibodies in a biological sample obtained from an animal. This is the
CC amino acid sequence of Staphylococcus aureus cbrA protein.
XX
SQ Sequence 330 AA;

Query Match 24.0%; Score 405; DB 7; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.4e-24;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY 8 LKLSVIGLFLVLIATAACGNNSSNSKESKGVIEIKHEGTTKVPKPKRVVVLEYS 67
DB 5 IKGL-VVTLAFILV-LAGCGSNKQSSNDKQKETSIKHMGTTIKGPKRVVTLVQG 62
QY 68 FVDALVALDVKPVGADDDNKKRIIKPLRDKIKGYTSVGRKQPNLEEISKLKPLIITAD 127
DB 63 ATDVAVSLGVKPVGAVESWTQPKFVYIKNDLKDTKIVGQEPAPNLEEISKLKPLIIVAS 122
QY 128 NNRHGIYKDLNKAIPITELKSPFDGNYENIDAFKTIKALKEBEGKRLKEHDKKIEE 187
DB 123 KVRNEKVDQLSKIAPTV---STDVTVPFKP-DTTKLMGKALGKEAEADLLKKYDDKVA 178
QY 188 YKKEITMDNQKVLPAVAAKSGLLAHPSN-----SYVGQFQLGFKFEALSDDVTKGL 240

Db 179 FOK-----DAKAKYKDAWPLKASVVNFRADHTRIYAGGYAGEILNDLGFKR-----NKDL 228
QY 241 SKYLKG--PYLQMT--ETLSQVNPFRMFMITNKASSNEPSL-----KELEKDPVWKKLN 292
DB 229 QKQVDNGKDIITQLTSKGSIPLMNADHIFVVKSDPNADKDALVVKTESEWTSKKEWNLDA 288
QY 293 VKNQRV--DILDRDLWARSRLISSEEMAKELVELSKDKKDNK 335
DB 289 VKNNQVSDDLDEITWNLGGYKSSKLIDDLVE--KLNIEKQSK 330

RESULT 30
ABU17675
ID ABU17675 standard; protein; 324 AA.
XX
AC ABU17675;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #3202.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bacillus anthracis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA21545.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 45599; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway; (8)
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
```


CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 324 AA;

Query Match 23.1%; Score 391; DB 6; Length 324;
 Best Local Similarity 34.4%; Pred. No. 1.9e-23;
 Matches 116; Conservative 59; Mismatches 122; Indels 40; Gaps 13;
 QY 9 KILSVIGLLFVLITAAAGNNSSNSSKESKD---GVEIKHEEGTTKVPKPKRVVVL 64
 DB 3 KILSI---FIVFLFVGCQQKEKETKADNKNQAITIKHAGETKLDKPAKKVVVL 58
 QY 65 EYSFVDALVALDVKPVGIADNDKNRII---KPLRDKIGKYTSVGTQKQNLBEISKLK 120
 DB 59 EWWYSEDLALGVQPVGNADIKYNNKWTYTKPSKDVV---DVGTRQPNLEISLKL 114
 QY 121 PDLIADNNRHGIYKDLNKIAPTIEL---KSPGDYNNENIDAKFTISKALGKKEEGKKR 177
 DB 115 PDLIITTSFRGKAIKNELEQIAPTVMFDPSTSNNDHFAEMTETFKQIAKAVGKEEGKKV 174
 QY 178 LEEHDKKIEEYKKEI---TWDRKQKVLPAVAKS---GLLAHPSNSYVQGFLSQLGFK 229
 DB 175 LADMDKAPADAKAIEKADLKDKNIAQAFTAKNVPTFRILT--DNSLALQVTKKLG-- 230
 QY 230 EALSDDVTGKLSKYLKGPYLQMTETLSQVNPFR-MFTMTNKASSNEPSLKELEKDPVWK 288
 DB 231 --LINTPEAGKSE--PDGFKTTVESLQVDSNFYIIVADEDNIDFTQLK--GNPAWE 283
 QY 289 KLNAXNQORVDILDRDLWARSGLISSEEMAKELVEL 325
 DB 284 ELKPKKENMYKLKGDWTFIG-GPESATSLATQVADV 319

RESULT 31

ABG08851
 ID ABG08851 standard; protein; 1132 AA.

XX AC ABG08851;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8942.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PP 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI, 2001-639362/73.

XX DR N-P8DB; AAS73038.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 39210; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1132 AA;

Query Match 22.2%; Score 375; DB 4; Length 1132;
 Best Local Similarity 32.5%; Pred. No. 2e-21;

Matches 96; Conservative 52; Mismatches 115; Indels 32; Gaps 5;

QY 45 IKHEEGTTPKPKRVVVLVLEYSFVDALVALDVKPVGIADNDKNRIIKPLRDKIGKYTS 104

DB 24 VQDEHGTILTLEKTPQRIIVLELSFADALAAVDVIPGIADNDKAKRILPEVRAHLKPQWS 83

QY 105 VGTQKQNLBEISKLKPDLIADNNRHGIYKDLNKIAPTIELKSPGDYNNENIDAKFTI 164

DB 84 VGTQKQNLBEISKLKPDLIADNNRHGIYKDLNKIAPTIELKSPGDYNNENIDAKFTI 143

QY 165 SKALGKEEGKKRLEEHDKKIEEYKKEITMDKQKVLPAVAKSGLLAHPSNSYVQGLS 224

DB 144 GEMVKKREMQARLEQHKERMAQWASQ--LPKGTVAFGTSREQQFNLHTQETWTGSLA 201

QY 225 QLGFKEALSDVDTGKLSKYLKGPYLQMTETLSQVNPFRMFTMTNKASSNEPSLKELEKD 284

DB 202 SLGL-----NVPAAAGASNPSPGPGASCPCRSIHAIWTSITPCL----- 245

QY 285 PVWKKLNAVKNQVRDIL-----DRDLWARSGLISSE-----EMAKELVELSK 327

DB 246 --WTPGSVIRCRPVGVLKMTDEAGEDAKLVAVPHSKLSKEYDHIKDVNDLPPELLK 298

RESULT 32

ABG13148

ID ABG13148 standard; protein; 932 AA.

XX AC ABG13148;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #13139.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

DT 26-SEP-2001 (first entry)
 DE C glutamicum protein fragment SEQ ID NO: 3875.
 XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 XX Corynebacterium glutamicum.
 OS EP1108790-A2.
 PN 20-JUN-2001.
 XX 18-DEC-2000; 2000EP-00127688.
 XX 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI; 2001-376931/40.
 DR N-PSDB; AAH65340.
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX Claim 17; SEQ ID NO 3875; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium, and identifying a homologue of a gene derived from
 CC corynebacterium. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX Sequence 315 AA;
 SQ
 Query Match 19.6%; Score 331; DB 4; Length 315;
 Best Local Similarity 28.6%; Pred. No. 1.5e-16;
 Matches 90; Conservative 65; Mismatches 140; Indels 20; Gaps 7;
 QY 13 VIGLFLVLIATAACGNNSSNS-SKESKDGVEIKHEGTTKPKRKRVLVLSYFVDA 71
 DB 15 VAPLGLVALLSSCSTSSDSIQPEVASTGYSVEHANGTTEIPETPRVVVDPSHLDA 74
 QY 72 LVALDVKPVGADNKNRIKPKRDKIGTKYTSVGRKQPNLEISKLPDLIIADNNRH 131
 DB 75 LLALGITPVGATESGSENGFPAYLADELKLTSEPLNLEKIAALDPPDLIIIGAKVRH 134
 QY 132 KGIYKDLNKAIPTELKSPDGDYENIDAPKTIISKALGKEEGKRLKEEHDKTIEYKKE 191
 DB 135 EAIYDQSDIAPTVMSESGSTGNWEQAE---ITAAVKNKSDMKLSDLDTRATELGE 191
 QY 192 ITMDKNQKVLPAVAAKSGLLAHPSNYSVVGQFLSQGFKEALSDVDTKLSKYLKGPYLQM 251
 DB 192 IGADQQTASVMVRFTDNFRLVGP-ETFGSVLEQVGP-----DLGERDNEY---SMWEL 242
 QY 252 NTEFLSQVNPMP--IMTNKASNEPSLKELEKDPVWKUNAVNQVRDILDRDLWARS 309
 DB 243 SSFNGDGLGILFIYTPGSPATYTPKISEL-----WVDSPAVRQGRTYEFEDETWVG 297
 QY 310 RGLISSEMAKELVE 324

DB 298 IGVLGANEILDDLEE 312
 RESULT 35
 ABU25343
 ID ABU25343 standard; protein; 321 AA.
 XX AC ABU25343;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #10870.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Clostridium difficile.
 OS WO200277183-A2.
 PN 03-OCT-2002.
 PD 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA29213.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 53267; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 321 AA;

Query Match 19.3%; Score 327; DB 6; Length 321;
Best Local Similarity 30.9%; Pred. No. 3.2e-18;
Matches 99; Conservative 55; Mismatches 136; Indels 30; Gaps 12;

QY 11 LSVIGLLFVLIATAAGNNSSSKSSKSGVGEIKHEEGTTKVPKPKRVVVLVSFVD 70
DB 11 LAILGLVTVF---ALGGSKNKSSTSENNITIKITHNLGETDVKLNKAKVVPFYSALD 67

QY 71 ALVALDVKP--VGIADNNKNNRIIKPLRDKIGKTSVGTGRKOPNLEESIKLKPDLIIADN 128
DB 68 TWDALGVAENLVGLPKASLPASLEK-YKDK--KYTDLGLGLPEPDLEGIKSANPDLIIN-N 123

QY 129 NRHKGIYKDLNKIAPTIELKSPDGYNE----NIDAFYTSIKALKEEGKRLSEHDKK 184
DB 124 GRQEDFYEQLSKIAPTISTSKDKKYLESVKNID---KIAKIFGVEEKANQEFSEKIEKK 180

QY 185 IEYKKEITMDKNQKVLPAVAAGSLLAHPSNYSVGGFQLSGLGKEALSDVDVTKGLSKYL 244
DB 181 IETLNKKVT-DKNLNAITIMWNEGNSLVGEESESRFSILYNSFGP-----ENKDKNIKESS 234

QY 245 KGPYLQMTETLSQVNPFRMFTM-----TNKASSNEPSLKELEKDPVWKKLNNAVKNQVD 299
DB 235 HQQ--NITPEYIAKQNPWFVDRGIATGSDVKESSTAKSVLNNDIISKMDAYKNDNII 292

QY 300 ILDRDLW-ARSRGLISSEEM 318
DB 293 YLDSPTWYVNDGGLTSLNKM 312

RESULT 36
ABU18359
ID ABU18359 standard; protein; 314 AA.
XX
AC ABU18359;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #3886.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bacillus anthracis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA22229.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 46283; 1766pp; English.
PS

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 314 AA;

Query Match 18.5%; Score 313; DB 6; Length 314;
Best Local Similarity 29.9%; Pred. No. 4.3e-17;
Matches 99; Conservative 65; Mismatches 135; Indels 32; Gaps 14;

QY 9 KILSVIGLLFVLIATAAGNNSSSKSSKSGVGEIKHEEGTTKVPKPKRVVVLVSF 68
DB 4 KLFILFTMLVLSIVGSSSQESKAKEQPKTKV-VKHAKGEATIPVNPKEIVDLSSG- 61

QY 69 VDIALVDVDPVGIADNNKNNRIIKPLRDKIGKTSVGTGRKOP--NLBEISIKLKPDLIIA 126
DB 62 TEELLLGLHKKPVGT-A-NYKDKIQKHLTEKLDGKAVGWYAPKVDLEAVTALKPDLLIL 120

QY 127 DNNRHGIYKDLNKIAPTIELKSPDGYNEIDAFYTSIKALKEEGKRLSEHDKKIE 186
DB 121 -NNRQLKIYDQLEKVAPTAVTLETNLEWR---GKFEVGEKLFDESKKADKWIADYDKKAD 176

QY 187 EYKKEITMDKNQKVLPAVAAGSLLAHPSNYSV--GQFLSQLGFKKALSDVDV---TKGLS 241
DB 177 SLSSKKIKKTKDDSPMFVAVT-----PQNFVYGSF-----GYGDIIINDLKLPAKGTGD 226

QY 242 KYLKGPYLQMTETLSQVNPFRMFTMNTKASSNEPSLKELEKDPVWKKLNNAVKNQVDIL 301
DB 227 --LKQTMQAQVLSLEGLVAFQPDQMFIV-NFGGEADKVVYEDYKNKSAVWKNKAVNNHVEV 283

QY 302 DRDLW-ARSRGLISS-----EEMAKELVELSK 327
DB 284 SNEVENTKAFNPIGKMDLIDIEIAKEILAKNK 314

RESULT 37
ADC94889
ID ADC94889 standard; protein; 352 AA.
XX
AC ADC94889;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 4516.

XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.
XX Enterococcus faecium.
XX US6583275-B1.
XX 24-JUN-2003.
XX 30-JUN-1998; 98US-00107532.
XX 02-JUL-1997; 97US-0051571P.
XX 14-MAY-1998; 98US-0085598P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm LA, Bush D;
XX WPI; 2003-799836/75.
XX N-PSDB; ADC91235.
XX New isolated nucleic acid derived from Enterococcus faecium encoding an
XX Enterococcus faecium polypeptide useful for detection, prevention and
XX treatment of a pathological condition resulting from a bacterial
XX infection.
XX Example 1; SEQ ID NO 4516; 243pp; English.
XX The invention relates to an isolated nucleic acid derived from
XX Enterococcus faecium encoding an Enterococcus faecium polypeptide having
XX one of 10 fully defined sequences given in the (or comprising 40
XX sequential nucleotides chosen from any of the nucleic acids, its
XX complement or sequences hybridising to it). Also included are a
XX recombinant vector comprising the nucleic acid operably linked to
XX a transcription regulatory element, a cell comprising the vector and a
XX single-stranded probe comprising the nucleic acid. The nucleic acids are
XX chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX The nucleic acid is useful for diagnosing pathological conditions
XX resulting from E. faecium bacterial infection (e.g. urinary tract
XX infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
XX infection) and for screening drugs such as agonists and antagonists. The
XX nucleic acid is useful for recombinant production of Candida albicans -
XX derived peptides or antisense polypeptides. Pharmaceutical compositions
XX and vaccines containing the nucleic acid are useful for preventing or
XX treating Enterococcus faecium infections. The present sequence represents
XX one if the disclosed E. faecium proteins.
XX Sequence 352 AA;
SQ
Query Match 18.2%; Score 307.5; DB 7; Length 352;
Best Local Similarity 31.3%; Pred. NO. 1.4e-16;
Matches 107; Conservative 65; Mismatches 123; Indels 47; Gaps 15;
QY 5 VRGKILSVIGLLFLIATAAG--NNSSNSSKSS---KDGVKHEEG-TTKVPKHP 58
DB 40 VVGILLASFG-----AACGNSNFTSQDTKRSQTQTTLTITDSNGDQIEVPNPP 92
QY 59 KRVVLEYSFYDALVALDV--KPGVIADNDKKNRIKPLRDKIGKYSVGRKQPNLEEI 116
DB 93 EKVVVFNGSLDTMDALGVGKVVGAATSS-----LPEYLSYKKVESAGGIKEPDLEKI 147
QY 117 SKLXPDLIIADNNRHKGIVKDLNKIAPTEILKSPFG-----DYNENIDAFKTIKALGKE 171
DB 148 NQLQPDLLII-SGRQRDPQKLSQTAPITFL-SLDKPNWPSFQONVTA---LGEIFGKQ 202
QY 172 BEGKRLKEEDHKTEEVKKEITMDKNQKVPAAKSGLLAHPNSVYVGQFLSQLGPKEA 231
DB 203 BEAKTQLEELSSALDQTKKAAEA-TDKXALVTLVNEGQLSAYGSGSRFGFTHDLFGFQA 261
QY 232 LSDP-----VTXGLSKYLKPGYLOMNTETLSQVNPERFIM-TNKASSNEPSLKEKDPV 286
DB 262 --DQIEASTHGQS-----VSVEYVLEKNPDILPVVDRTKAIGGDDSKDDISANEL 310

QY 287 WKKLNAVQNRVDILDRDLWARSGRLISSEMAKVELSKK 328
DB 311 VAQTNAGNQQIISLEPDVWYLSGGUESMKIMIEDVNAFK 352
RESULT 38
ABU17461
ID ABU17461 standard; protein; 314 AA.
XX AC ABU17461;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #2988.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Bacillus anthracis.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA21331.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 45385; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of

KW antibacterial; gene therapy.
OS Staphylococcus epidermidis.
XX US6380370-B1.
XX 30-APR-2002.
XX
XX 13-AUG-1998; 98US-00134001.
XX
XX 14-AUG-1997; 97US-0055779P.
XX 08-NOV-1997; 97US-0064964P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX WPI; 2002-381255/41.
XX N-PSDB; ABN93192.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
XX polypeptide, useful for diagnosing and treating bacterial infections.
XX
XX Disclosure; SEQ ID NO 5492; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX antibacterial activity and can be used in gene therapy. The sequences can
XX also be used in the diagnosis and treatment of bacterial infections,
XX particularly S. epidermidis infections. The sequences can be used to
XX screen for compounds able to interfere with the S. epidermidis life cycle
XX or inhibit S. epidermidis infection. N.B. The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the USPTO web site
SQ Sequence 351 AA;

Query Match 17.5%; Score 296.5; DB 5; Length 351;
Best Local Similarity 28.4%; Pred. No. 1.1e-15;
Matches 104; Conservative 70; Mismatches 131; Indels 61; Gaps 16;

Qy 1 GVBSVRGLKILSLVIGLLFLVLIATAACGNNSSNSSKES-SKDGVETKHE----- 48
Db 1 GVAHMKTVLFLLSLVLTACSNSSNNNSISKKNSDSKETVTIKNSPESKKNNGS 60

Qy 49 -----EGTTKVPKPRVWVLEYSFVDALVALDVKPVGIADDNK-----KNRIKPLRD 97
Db 61 DKKKISNTVEVPKPKNAVLDYGDVLDL-----KELGVADKVKGLPKGENNQSLPKFLD 115

Qy 98 --KIGKTSVQTRKQPNLEISKLKPDLI-IADNNRHKGIYKOLNKTAPIELKSPGDY 154
Db 116 EFKDDKYINTGNLKEVNFVKVAKPDVIFISGRGTANQKNLDDEFKKAAPKAKV-VYVGTS 174

Qy 155 NENI--DAFK--TISKALGKEEGKRLSEHDKKIEYKKEITMDKNQKVLPAVAAKSG 209
Db 175 DDNLIKDMKNNTENIGKIYDKEDAKKINKDLDRKISDMKDK-TDFNKKVWYLLVNEGE 233

Qy 210 LLAH-PSNSYVGQFLSQLGPKFKAALSDVTKGLSKYLKGPYLQ-MNTETLSQVNP----- 261
Db 234 LSTFGPGRFGGLVFDTLGFKPA-----DKKVSXSPHGQININNEYINKQNPVDILAM 285

Qy 262 ERMFIMTNKASSNPSLKELEKOPVWKKLNKQVQRVDILDRDLWARSGLISSEEMAKE 321
Db 286 DRGSVVGKATTN-----QVLKNKVIKNVKAHNIYELDPKLWYFSSG--SSTTTIKQ 338

Qy 322 LVELSK 327
Db 339 IDELNE 344